DUKE MATHEMATICAL JOURNAL

VOLUME 3
1937

DUKE UNIVERSITY PRESS
DURHAM, N. C.

CONTENTS

VOLUME 3, 1937

SUMS OF SQUARES OF POLYNOMIALS

By Leonard Carlitz

1. **Introduction.** In this note we determine the number of representations of 0 as the sum of an arbitrary number of squares of polynomials in a single indeterminate with coefficients in a fixed Galois field $GF(p^n)$, $p > 2$. More accurately, if $\alpha_1, \cdots, \alpha_t$ are t non-zero elements of $GF(p^n)$, $\alpha_1 + \cdots + \alpha_t = 0$, we determine the number of solutions of

$$0 = \alpha_1 Y_1^2 + \cdots + \alpha_t Y_t^2 \tag{1.1}$$

in primary[1] polynomials Y_j each of degree k, an assigned positive integer. We denote the number of solutions of (1.1) by

$$N_t(0) = N_t^k(0).$$

The more general equation

$$\alpha G = \alpha_1 Y_1^2 + \cdots + \alpha_t Y_t^2, \tag{1.2}$$

where $\alpha G \neq 0$, G of degree $\leqq 2k$, has been treated in two papers, one on the case t even, the other on the case t odd.[2] In the latter paper a formula for $N_{2s}(0)$ appeared incidentally. We shall derive this formula anew by the simpler and direct method used in the paper on t even.

To evaluate $N_{2s+1}(0)$, we make use of a known formula for $N_{2s}(G)$, the number of solutions of (1.2) for $t = 2s$. Applying this formula, we first evaluate the sums

$$\sum_G \frac{N_{2s}(G)}{|G|^w}, \qquad \sum_G \frac{N_{2s}(G^2)}{|G|^w},$$

extended over all primary G; the latter sum leads at once to the determination of $N_{2s+1}(0)$.

2. **Determination of** $N_{2s}(0)$. In equation (1.2), let $t = 2s$, $\alpha = \alpha_1 + \cdots + \alpha_{2s} \neq 0$, so that G is of degree $2k$. Assume further

$$\gamma_i = \alpha_{2i-1} + \alpha_{2i} \neq 0 \qquad (i = 1, \cdots, s). \tag{2.1}$$

Received October 27, 1936.

[1] A polynomial is *primary* if the coefficient of the highest power of the indeterminate is the unit element of the Galois field. The capitals A, B, E, G, M, U, V, Y will denote primary polynomials.

[2] The even case in Transactions of the American Mathematical Society, vol. 35 (1933), pp. 397–410; the odd case in this Journal, vol. 1 (1935), pp. 298–315. These papers will be cited as I and II, respectively.

Then by Theorem 4 of paper I, the number of solutions of (1.2) is determined by

$$N_{2s}(G) = \left(1 - \frac{\chi}{p^{n(s-1)}}\right) \sum_{M \mid G}^{m>k} \chi^m \mid M \mid^{s-1} + \sum_{M \mid G}^{m=k} \chi^m \mid M \mid^{s-1}. \tag{2.2}$$

Here $\chi = +1$ or -1 according as

$$(-1)^s \alpha_1 \alpha_2 \cdots \alpha_{2s} \tag{2.3}$$

is a square or a non-square in $GF(p^n)$; m is the degree of M, $\mid M \mid = p^{nm}$; the summations are extended over all M dividing G of degree $m > k$ and $m = k$, respectively. If now we denote the right member of (2.2) by $\Lambda_{s-1}(G, \chi)$, we shall prove the following formula:

$$\begin{aligned} &\sum_{\deg G=2k} \Lambda_s(G, \chi_1)\Lambda_t(G, \chi_2) \\ &\qquad = \left(1 - \frac{\chi}{p^{n(s+t+1)}}\right) \sum_{u=0}^{k-1} \chi^u p^{n(2k-u)(s+t+1)} p^{nu} + \chi^k p^{nk(s+t+2)}, \end{aligned} \tag{2.4}$$

where $\chi = \chi_1 \chi_2$, and the summation in the left member is over all primary G of degree $2k$.

The proof of this formula is similar to the proof of Theorem 2 of paper I, and therefore only a sketch of it will be given. By means of (2.2), the left member of (2.4) may be put in the form

$$\begin{aligned} &\left\{\left(1 - \frac{\chi_1}{p^{ns}}\right)\left(1 - \frac{\chi_2}{p^{nt}}\right) \sum_{\substack{u>k \\ v>k}} + \left(1 - \frac{\chi_1}{p^{ns}}\right) \sum_{\substack{u>k \\ v=k}} \right. \\ &\qquad\qquad \left. + \left(1 - \frac{\chi_2}{p^{nt}}\right) \sum_{\substack{u=k \\ v>k}} + \sum_{\substack{u=k \\ v=k}} \right\} \chi_1^u \chi_2^u \mid U \mid^s \mid V \mid^t, \end{aligned}$$

where each summation is over all A, B, U, V of degree a, b, u, v, respectively, such that

$$AU = BV, \qquad a + u = b + v = 2k,$$

and in addition satisfies the conditions under each Σ. Call the sums Σ_1, Σ_2, Σ_3, Σ_4, respectively. Then

$$\begin{aligned} \sum\nolimits_1 &= \sum_{\substack{AU=BV \\ u,v>k}} \chi_1^u \chi_2^v \mid U \mid^s \mid V \mid^t \\ &= p^{2nk(s+t)} \sum_{\substack{AU=BV \\ a,b<k}} \chi_1^a \chi_2^b \mid A \mid^{-s} \mid B \mid^{-t} \\ &= p^{2nk(s+t)} \sum_{\deg M<k} \chi^m \mid M \mid^{-s-t} S_M, \end{aligned}$$

where

$$S_M = \sum_{\substack{(A,B)=1 \\ a,b<k-m}} \chi_1^a \chi_2^b \mid A \mid^{-s} \mid B \mid^{-t} \sum_{\substack{AU=BV \\ a+u=b+v=2k-m}} 1. \tag{2.5}$$

Here the outer sum is over all A, B of degree $a, b < k - m$ such that $(A, B) = 1$, while the inner sum is over all U, V of degree $2k - m - a$, $2k - m - b$, respectively, for which $AU = BV$. Since $(A, B) = 1$, it follows that $A \mid V$ and $B \mid U$. For the moment put $V = AE$, $U = BE$, so that $\deg E = 2k - m - a - b$. Thus for fixed A, B there are $| E |$ choices of U, V satisfying the conditions mentioned. In other words, the inner sum in (2.5) $= p^{n(2k-m-a-b)}$, and therefore

$$S_M = p^{n(2k-m)} \sum_{a,b<k-m} \chi_1^a \chi_2^b p^{-na(s+1)-nb(t+1)} \psi(a, b), \tag{2.6}$$

where[3] $\psi(a, b)$ denotes the number of pairs A, B of degree a, b, respectively, such that $(A, B) = 1$. The sum in (2.6) is easily evaluated but this need not be repeated here. It is sufficient to notice that comparing this point in the present proof with the corresponding point[4] in the proof of Theorem 2 of paper I, we are at once able to conclude that the left member of (2.4)

$$\begin{aligned} &= p^{2nk(s+t+1)} \sum_{\deg M=k} \chi^m \,|\, M \,|^{-s-t-1} \\ &+ \left(1 - \frac{\chi}{p^{n(s+t+1)}}\right) p^{2nk(s+t+1)} \sum_{\deg M<k} \chi^m \,|\, M \,|^{-s-t-1} \\ &= \chi^k p^{nk(s+t+2)} + \left(1 - \frac{\chi}{p^{n(s+t+1)}}\right) \sum_{m=0}^{k-1} \chi^m p^{n(2k-m)(s+t+1)} p^{nm}. \end{aligned}$$

We have therefore proved (2.4).

It is now easy to evaluate $N_{2s}(0)$. Since the conditions (2.1) are assumed to hold, we have in particular $\alpha = \alpha_1 + \cdots + \alpha_{2s-2} \neq 0$. Then clearly all solutions of

$$0 = \alpha_1 Y_1^2 + \cdots + \alpha_{2s} Y_{2s}^2 \qquad (\deg Y_i = k)$$

may be obtained by pairing the solutions of

$$\alpha G = \alpha_1 Y^2 + \cdots + \alpha_{2s-2} Y_{2s-2}^2 \qquad (\deg Y_i = k)$$

with those of

$$-\alpha G = \alpha_{2s-1} Y_{2s-1}^2 + \alpha_{2s} Y_{2s}^2 \qquad (\deg Y_i = k)$$

in all possible ways, and allowing G to range over all polynomials of degree $2k$. We define χ_1 as $+1$ or -1 according as $(-1)^{s-1}\alpha_1\alpha_2 \cdots \alpha_{2s-2}$ is or is not a square in $GF(p^n)$; χ_2 is defined in the same way with respect to $-\alpha_{2s-1}\alpha_{2s}$. Then, by the remark just made, it follows that

$$N_{2s}(0) = \sum_{\deg G=2k} N_{2s-2}(G) N_2(G),$$

and by (2.2), the right member

$$= \sum_{\deg G=2k} \Lambda_{s-2}(G, \chi_1) \Lambda_0(G, \chi_2).$$

[3] See I, p. 399.

[4] I, p. 405 bottom; cf. also proof of I. Theorem 1.

We may now apply (2.4), and we have at once

$$N_{2s}(0) = p^{nks} + \left(1 - \frac{\chi}{p^{n(s-1)}}\right) \sum_{m=0}^{k-1} \chi^m p^{n(2k-m)(s-1)} p^{nm}. \tag{2.7}$$

Thus we have the following

THEOREM A. *If $\alpha_1, \cdots, \alpha_{2s}$ are $2s$ non-zero elements of $GF(p^n)$, such that $\alpha_1 + \cdots + \alpha_{2s} = 0$, while*

$$\alpha_{2i-1} + \alpha_{2i} \neq 0 \qquad (i = 1, \cdots, s),$$

then the number of solutions of $\alpha_1 Y_1^2 + \cdots + \alpha_{2s} Y_{2s}^2 = 0$ in primary Y_j of degree k is furnished by (2.7) *together with* (2.3).

3. **The sum $\Sigma N_{2s}(G) \mid G \mid^{-w}$.** In (2.2) G is assumed to be of even degree. However, if we rewrite the equation in the following form

$$N_{2s}(G) = \left(1 - \frac{\chi}{p^{n(s-1)}}\right) \sum_{G=MU}^{m>u} \chi^m \mid M \mid^{s-1} + \sum_{G=MU}^{m=u} \chi^m \mid M \mid^{s-1}, \tag{3.1}$$

then $N_{2s}(G)$ is defined for all G (for G of odd degree the second sum on the right is vacuous). We may now consider the sum $\Sigma N_{2s}(G) \mid G \mid^{-w}$ taken over *all* primary G. Thus substituting from (3.1), we have

$$\sum_G \frac{N_{2s}(G)}{\mid G \mid^w} = \sum_{m=u} \frac{\chi^m \mid M \mid^{s-1}}{\mid MU \mid^w} + (1 - \chi p^{n(1-s)}) \sum_{m>u} \frac{\chi^m \mid M \mid^{s-1}}{\mid MU \mid^w}, \tag{3.2}$$

where the sums on the right are over all M, U of degree m, u, respectively, satisfying the conditions indicated. From this it is clear that the right member of (3.2) may be reduced to

$$\sum_{m=0}^{\infty} \chi^m p^{nm(s+1-2w)} + (1 - \chi p^{n(1-s)}) \sum_{u=0}^{\infty} p^{nu(1-w)} \sum_{m=u+1}^{\infty} \chi^m p^{nm(s-w)}$$

$$= \frac{1}{1 - \chi p^{n(s+1-2w)}} + (1 - \chi p^{n(1-s)}) \sum_{u=0}^{\infty} \chi^u p^{nu(s+1-2w)} \frac{\chi p^{n(s-w)}}{1 - \chi p^{n(s-w)}}$$

$$= \frac{1}{1 - \chi p^{n(s+1-2w)}} + \frac{1 - \chi p^{n(1-s)}}{1 - \chi p^{n(s+1-2w)}} \frac{\chi p^{n(s-w)}}{1 - \chi p^{n(s-w)}},$$

from which it follows that

$$\sum_G \frac{N_{2s}(G)}{\mid G \mid^w} = \frac{1 - p^{n(1-w)}}{(1 - \chi p^{n(s+1-2w)})(1 - \chi p^{n(s-w)})}. \tag{3.3}$$

If we split the right member into partial fractions, (3.3) becomes

$$\sum_G \frac{N_{2s}(G)}{\mid G \mid^w} = \frac{1}{1 - \chi p^{n(s-w)}} - \frac{p^{n(1-w)}}{1 - \chi p^{n(s+1-2w)}}. \tag{3.4}$$

From this it is easy to evaluate $\sum N_{2s}(G)$, summed over all G of fixed degree. For even degree we have the following simple formula:

$$\sum_{\deg G=2k} N_{2s}(G) = p^{2nks}, \tag{3.5}$$

while for odd degree,

$$\sum_{\deg G=2k+1} N_{2s}(G) = \chi p^{n(2k+1)s} - \chi p^{nk(s+1)} p^{n}. \tag{3.6}$$

The sum formula for even degree (3.5) may be derived in a slightly more direct manner. We consider the sum $\sum N_{2s}(G) \mid G \mid^{-w}$ taken over G of even degree only, so that $N_{2s}(G)$ has its original meaning. Then as above,

$$\begin{aligned}
\sum_{\deg G \text{ even}} \frac{N_{2s}(G)}{|G|^w} &= \sum_{m=u} \frac{\chi^m |M|^{s-1}}{|MU|^w} + (1 - \chi p^{n(1-s)}) \sum_{\substack{m>u \\ m-u \text{ even}}} \frac{\chi^m |M|^{s-1}}{|MU|^w} \\
&= \sum_{m=0}^{\infty} \chi^m p^{nm(s+1-2w)} + (1 - \chi p^{n(1-s)}) \sum_{u=0}^{\infty} p^{nu(1-w)} \sum_{m-u=2,4,\cdots} \chi^m p^{nm(s-w)} \\
&= \frac{1}{1 - \chi p^{n(s+1-2w)}} + (1 - \chi p^{n(1-s)}) \sum_{u=0}^{\infty} \chi^u p^{nu(s+1-2w)} \frac{p^{2n(s-w)}}{1 - p^{2n(s-w)}} \\
&= \frac{1}{1 - \chi p^{n(s+1-2w)}} + \frac{1 - \chi p^{n(1-s)}}{1 - \chi p^{n(s+1-2w)}} \frac{p^{2n(s-w)}}{1 - p^{2n(s-w)}},
\end{aligned}$$

from which we have at once

$$\sum_{\deg G \text{ even}} \frac{N_{2s}(G)}{|G|^w} = \frac{1}{1 - p^{2n(s-w)}}; \tag{3.7}$$

this agrees with (3.5).

Comparing (3.7) with (3.4), we have

$$\sum_{\deg G \text{ odd}} \frac{N_{2s}(G)}{|G|^w} = \frac{\chi p^{n(s-w)}}{1 - p^{2n(s-w)}} - \frac{p^{n(1-w)}}{1 - \chi p^{n(s+1-2w)}},$$

which is equivalent to (3.6).

4. **The sum** $\sum N_{2s}(G^2) \mid G \mid^{-w}$. The evaluation of this sum is somewhat more elaborate than that of the sum (3.2). If we put $G^2 = M_0 U_0$, and write $(M_0, U_0) = D$, then clearly M_0/D and U_0/D are squares; we may therefore suppose that

$$G^2 = DM^2 \cdot DU^2, \qquad G = DMU, \qquad (M, U) = 1.$$

Then substituting from (2.2), we now have

$$\sum_G \frac{N_{2s}(G^2)}{|G|^w} = (1 - \chi p^{n(1-s)}) \sum_{\substack{D,M,U \\ m>u}}' \chi^d \frac{|DM^2|^{s-1}}{|DMU|^w} + \sum_{\substack{D,M,U \\ m=u}}' \chi^d \frac{|DM^2|^{s-1}}{|DMU|^w}, \tag{4.1}$$

where the sum on the left is over all G, while the sums on the right are over all D, M, U of respective degree d, m, u satisfying the condition under each $\sum'$, and in addition $(M, U) = 1$. We may evidently put the right member of (4.1) in the form

$$\sum_{d=0}^{\infty} \chi^d p^{nd(s-w)} T = \frac{T}{1 - \chi p^{n(s-w)}}, \tag{4.2}$$

where

$$T = \sum_{\substack{(M,U)=1 \\ m=u}} | M |^{2s-2-w} | U |^{-w} + (1 - \chi p^{n(1-s)}) \sum_{\substack{(M,U)=1 \\ m>u}} | M |^{2s-2-w} | U |^{-w}. \tag{4.3}$$

As for the first sum in (4.3), for $\psi(a, b)$ as defined in (2.6), it is clear that

$$\sum_{\substack{(M,U)=1 \\ m=u}} | M |^{2s-2-w} | U |^{-w} = \sum_{m=0}^{\infty} p^{nm(2s-2-2w)} \psi(m, m),$$

$$= 1 + (1 - p^{-n}) \sum_{1}^{\infty} p^{nm(2s-2w)}$$

$$= \frac{1 - p^{n(2s-1-2w)}}{1 - p^{n(2s-2w)}}. \tag{4.4}$$

The calculation of the second sum in (4.3) is somewhat longer. In the first place

$$\sum_{\substack{(M,U)=1 \\ m>u}} | M |^{2s-2-w} | U |^{-w} = \sum_{m>u} p^{nm(2s-2-w)} p^{-nuw} \psi(m, u).$$

The terms in which $u = 0$ contribute

$$\sum_{m=1}^{\infty} p^{nm(2s-2-w)} p^{nm} = \frac{p^{n(2s-1-w)}}{1 - p^{n(2s-1-w)}}. \tag{4.5}$$

For the remaining terms, we have

$$\sum_{m>u>0} p^{nm(2s-1-w)} p^{nu(1-w)} (1 - p^{-n})$$

$$= (1 - p^{-n}) \sum_{u=1}^{\infty} p^{nu(1-w)} \sum_{m=u+1}^{\infty} p^{nm(2s-1-w)}$$

$$= (1 - p^{-n}) \frac{p^{n(4s-1-3w)}}{(1 - p^{n(2s-2w)})(1 - p^{n(2s-1-w)})}, \tag{4.6}$$

by an easy calculation. Combining (4.4) and (4.5), we see that the second sum in (4.3)

$$= \frac{p^{n(2s-1-w)} - p^{n(4s-2-3w)}}{(1 - p^{n(2s-2w)})(1 - p^{n(2s-1-w)})}.$$

Substituting from (4.4) and (4.5) in (4.3), we may verify that

$$T = \frac{(1 - \chi p^{n(s-w)})(1 - p^{n(2s-1-2w)})}{(1 - p^{n(2s-2w)})(1 - p^{n(2s-1-w)})}.$$

Therefore, returning to (4.2), we see finally that

$$\sum_G \frac{N_{2s}(G^2)}{|G|^w} = \frac{1 - p^{n(2s-1-2w)}}{(1 - p^{n(2s-2w)})(1 - p^{n(2s-1-w)})}. \tag{4.7}$$

5. **Determination of** $N_{2s+1}(0)$. Making use of the formula just proved, we shall now evaluate $N_{2s+1}(0)$. Let $\alpha_1, \cdots, \alpha_{2s+1}$ be $2s + 1$ non-zero elements of $GF(p^n)$ such that $\alpha_1 + \cdots + \alpha_{2s+1} = 0$. We may assume them so numbered that the conditions (2.1) are satisfied,[5] and therefore $N_{2s}(G)$ applies and is determined by (2.2). Now the number of solutions of

$$0 = \alpha_1 Y_1^2 + \cdots + \alpha_{2s+1} Y_{2s+1}^2 \qquad (\deg Y_j = k)$$

is clearly the sum of the number of solutions of

$$-\alpha_{2s+1} G^2 = \alpha_1 Y^2 + \cdots + \alpha_{2s} Y_{2s}^2 \qquad (\deg Y_j = k),$$

extended over all G of degree k. In other words,

$$N_{2s+1}(0) = \sum_{\deg G = 2k} N_{2s}(G^2). \tag{5.1}$$

But the sum in the right member is precisely the coefficient of p^{-nkw} in the expansion of

$$\sum_G \frac{N_{2s}(G^2)}{|G|^w} = \sum_{k=0}^{\infty} \frac{1}{p^{nkw}} \sum_{\deg G = k} N_{2s}(G^2). \tag{5.2}$$

Comparing (5.2) with (4.7), it is clear that the quantity in question is the coefficient of p^{-nkw} in

$$(1 - p^{n(2s-1-2w)}) \sum_{i,j=0}^{\infty} p^{n(2is-2iw)} p^{n(2s-1)j-njw}$$

$$= (1 - p^{n(2s-1-2w)}) \sum_k \{p^{n(2s-1)k} + p^{n(2s-1)(k-2)} p^{2ns} + p^{n(2s-1)(k-4)} p^{4ns} + \cdots\} p^{-nkw}.$$

Therefore, finally

$$N_{2s+1}(0) = p^{n(2s-1)k} + (1 - p^{-n}) \sum_{0 < 2i \leq k} p^{n(2s-1)(k-2i)} p^{2nis}. \tag{5.3}$$

For $k < 2$, this reduces to the first term: $N_{2s+1}(0) = p^{n(2s-1)k}$. According to (5.3), $N_{2s+1}(0)$ depends only on s and k; in particular it is independent of the numbering of the α_j required at the beginning of this section. We may now state

THEOREM B. *If* $\alpha_1, \cdots, \alpha_{2s+1}$ *are* $2s + 1$ *non-zero elements of* $GF(p^n)$, *such that* $\alpha_1 + \cdots + \alpha_{2s+1} = 0$, *then the number of solutions of* $\alpha_1 Y_1^2 + \cdots + \alpha_{2s+1} Y_{2s+1}^2 = 0$ *in primary* Y_j *of degree* k *is furnished by* (5.3).

DUKE UNIVERSITY.

[5] See II, §3.

THE DEGREES OF THE IRREDUCIBLE COMPONENTS OF SIMPLY TRANSITIVE PERMUTATION GROUPS

By J. Sutherland Frame

1. In a study of certain hyperorthogonal groups,[1] there arose the problem of splitting into its r irreducible components a simply transitive permutation group G^* of degree n and order g, which, when written in matrix form, gave an isomorphic representation of the given abstract group as a group G of linear transformations. In any simply transitive permutation group, the subgroup leaving one symbol invariant will permute the remaining symbols in $\lambda = r - 1$ sets of transitivity of $k_1, k_2, \cdots, k_\lambda$ symbols respectively. Let the distinct irreducible components of the group G have the degrees $n_0 = 1, n_1, \cdots, n_{\lambda'}$, and note that $r = \lambda + 1$ is the sum of the squares of the multiplicities with which these occur in the reduction of G.[2] When the components are all distinct, and $\lambda' = \lambda$, there appears to be a simple relation between the product $K = 1 \cdot k_1 k_2 \cdots k_\lambda$ and the product $N = 1 \cdot n_1 n_2 \cdots n_\lambda$.

Conjectured Theorem I. *$n^{\lambda-1}K/N$ is an integer when the components of G are distinct, and this is a perfect square R^2 when the numbers k_i are distinct.*

We shall prove the theorem for all groups for which $\lambda \leqq 3$, (here λ' must equal λ), and for an infinite family of groups including all values of λ. When $\lambda = 1$, the group G^* is doubly transitive and $n_1 = k_1 = N = K = n - 1$, so the result is trivial. When $\lambda = 2$, our theorem gives us a diophantine equation, $nk_1k_2/n_1n_2 = R^2$, which, with $n_1 + n_2 = n - 1$, enables us to solve for the unknowns n_1 and n_2.

To illustrate the application of the theorem, before passing to the details of the proof, we take as an example the case of the hyperorthogonal groups, where the problem of this paper was suggested.[1] We have here a permutation group of degree Q_mQ_{m-1}/Q_2 (where $Q_m = q^m - (-1)^m$, $q = p^s$, p prime) which is known to have 3 irreducible components. We know also that $k_1 = q^{2m-3}$, $k_2 = q^2Q_{m-2}Q_{m-3}/Q_2$. Hence, $n_1n_2 = q^{2m-1}Q_mQ_{m-1}Q_{m-2}Q_{m-3}/Q_2^2R^2$, where R is an integer, and $n_1 + n_2 = n - 1 = (Q_mQ_{m-1} - Q_2)/Q_2$. The degree of n_1 or n_2 as a polynomial in q is $2m - 3$, that of the other being less; so the degree of R in q is at least $m - 2$. Since n_1n_2 is divisible by an odd power of q, and $n_1 + n_2$ is divisible by q^2, it follows that n_1, say, is divisible by q^2 but not q^3, and n_2 by q^3 or some higher odd power. R, not being divisible by q^{m-2}, must contain a factor

Received September 17, 1936.

[1] J. S. Frame, *Unitäre Matrizen in Galoisfeldern*, Commentarii Mathematici Helvetici, vol. 7 (1935), pp. 97, 98.

J. S. Frame, *The simple group of order 25920*, this Journal, vol. 2 (1936), p. 477.

[2] W. Burnside, *On the complete reduction of any transitive permutation group*, Proc. Lond. Math. Soc., ser. 2, vol. 3 (1905), p. 239.

whose square divides $Q_m Q_{m-1} Q_{m-2} Q_{m-3}/Q_2^2$. The only such factor is Q_1, or possibly Q_3 for certain values of m. The latter possibility does not work, and we find $R = q^{m-3}Q_1$, $n_1 = q^2 Q_m Q_{m-3}/Q_2 Q_1$, $n_2 = q^3 Q_{m-1} Q_{m-2}/Q_2 Q_1$. The degrees of the irreducible components are hereby uniquely determined from our formula.

In attacking the proof of the theorem, we shall first study in §2 the ring of matrices permutable with the group G, and then the secular equation, in terms of whose roots ρ_i the constants n_i may be expressed. In §3 we shall give the proof of the theorem for the cases $\lambda = 2$, $\lambda = 3$. In §4 we shall prove it for a certain infinite family of groups including all values of λ by solving explicitly for the constants n_i. A lemma on binomial coefficients will be proved in §5.

2. To each permutation $\binom{x_\nu}{x_{\nu'}}$ of the group G^*, there corresponds in G a matrix of degree n, $(\delta_{\alpha'\beta})$, where $\alpha, \beta = 1, 2, \cdots, n$, which may be thought of as transforming the space of the variables $x_1 \cdots x_n$. Let the subgroup G_1 of G leave x_1 invariant and permute the remaining variables in $\lambda = r - 1$ sets of transitivity T_t, of K_t variables respectively, $t = 1, \cdots, \lambda$. Then, setting $k_0 = 1$, we have

$$\sum_{t=0}^{\lambda} k_t = n. \tag{2.1}$$

Now if x_t be any variable from the set T_t, and we transform the subscripts of the product $\bar{x}_1 x_t$ by each permutation S of G^*, so that S transforms x_t into Sx_t, and $\bar{x}_t$ into $S\bar{x}_t = \overline{Sx_t}$, then the sum $(nk_t/g) \sum_{S \text{ in } G} (S\bar{x}_1)(Sx_t)$ is an invariant hermitian form H_t. Its non-vanishing coefficients are all 1, since the subgroup leaving x_1 and x_t both invariant is of index nk_t under G. Its matrix of coefficients V_t is one of the $\lambda + 1$ basis elements of the ring of matrices V which are permutable with each of the permutations of G.[3] The conjugate imaginary form $\bar{H}_t = H_{t'}$ is also invariant, and so the corresponding matrix $V_{t'}$, the transposed of V_t, is in the ring. In particular, H_0 is the unit hermitian form, and $V_0 = I$ the identity matrix. Furthermore, $W = \sum_{t=0}^{\lambda} V_t$ is a matrix consisting entirely of 1's.

Let $V = \sum_{i=0}^{\lambda} \alpha_i V_i$ be an arbitrary matrix of the ring, and let products be given by the formulas

$$V_i V_j = \sum_{s=0}^{\lambda} c_{ijs} V_s \tag{2.2}$$

$$V V_j = \sum_{s=0}^{\lambda} C_{js} V_s, \quad \text{where} \quad C_{js} = \sum_{i=0}^{\lambda} \alpha_i c_{ijs}. \tag{2.3}$$

[3] I. Schur, *Zur Theorie der einfach transitiven Permutationsgruppen*, Berlin, Sitzungsberichte (1933), pp. 598–623. Schur has developed several of the properties of this matrix ring, applying it particularly to the study of permutation groups having as a subgroup a regular group of the same degree. We assume no such subgroup here—in fact, none exists for the simple groups in which we are particularly interested.

Since $V_i W = k_i W$, we have

$$\sum_{j=0}^{\lambda} c_{ijs} = k_i. \tag{2.4}$$

In order to display the reducibility of G, we must change the variables in such a way that the invariant hermitian forms H_t become diagonal forms $S^{-1}H_tS$ on n_t variables respectively, $t = 0, 1, \cdots, \lambda$. The matrices V_t will then also be brought into diagonal form $S^{-1}V_tS$, and can be written as linear combinations of new basis elements M_i, each of which is merely a multiplication on a set of n_i of the new variables.

$$MM_i = \rho_i M_i, \qquad M = S^{-1}VS, \qquad \rho_i = \sum_{t=0}^{\lambda} \rho_{it}\alpha_t. \tag{2.5}$$

Since, when we choose the parameters α_t so that $M = M_j$, we have $M_jM_i = \rho_i M_i = 0$, but $M_jM_j = \rho_j M_j \neq 0$, no two multipliers ρ_i and ρ_j corresponding to different M_i and M_j can be identical in the α_t. Each multiplier ρ_i is a root of multiplicity n_i of $\det(V - \rho I) = 0$. More simply, it is one of the $\lambda + 1$ distinct roots of the secular equation

$$D(\rho) \equiv \det(C_{ij} - \rho\delta_{ij}) \equiv (\rho_0 - \rho)(\rho_1 - \rho) \cdots (\rho_\lambda - \rho) = 0. \tag{2.6}$$

By adding all the other rows to any particular row of the determinant $D(\rho)$, we readily verify by (2.3) and (2.4) that

$$\rho_0 = \sum_{i=0}^{\lambda} k_i\alpha_i = \sum_{i=0}^{\lambda} C_{ij} \tag{2.7}$$

is a root. In particular, setting $\alpha_i = 1$, we find that $\rho_0 = n$ is only a simple root of $\det(W - \rho I) = 0$, the other roots being all 0. Hence ρ_0 is the root of multiplicity $n_0 = 1$, and each of the other roots are functions only of the differences $\alpha_i - \alpha_0$. If we denote the trace of the matrix V^h by ns_h, then

$$ns_h = \sum_{i=0}^{\lambda} n_i\rho_i^h, \tag{2.8}$$

and we may solve for the constants n_t. We let Δ be the Vandermonde determinant of $\rho_0, \rho_1, \cdots, \rho_\lambda$, and Δ_t be the Vandermonde determinant of $\rho_0, \rho_1, \cdots, \rho_{t-1}, \rho_{t+1}, \cdots, \rho_\lambda$, noting that $(-1)^{t+1}D'(\rho_t) = \Delta/\Delta_t$, and obtain

$$\frac{n_t}{n}\Delta = \begin{vmatrix} 1 & \cdots & 1 & 1 & 1 & \cdots & 1 \\ \rho_0 & \cdots & \rho_{t-1} & s_1 & \rho_{t+1} & \cdots & \rho_\lambda \\ \rho_0^2 & \cdots & \rho_{t-1}^2 & s_2 & \rho_{t+1}^2 & \cdots & \rho_\lambda^2 \\ \cdot & \cdots & \cdot & \cdot & \cdot & \cdots & \cdot \\ \rho_0^\lambda & \cdots & \rho_{t-1}^\lambda & s_\lambda & \rho_{t+1}^\lambda & \cdots & \rho_\lambda^\lambda \end{vmatrix} = F_t\Delta_t, \tag{2.9}$$

where F_t is the polynomial obtained from $(-1)^{t+1}D'(\rho_t)$ by replacing each power ρ_t^h of ρ_t by the corresponding s_h. Since $n_0 = 1$, we have

$$F_0 = \Delta/n\Delta_0 = -D'(\rho_0)/n, \tag{2.10}$$

$$F_t = n_t\Delta/n\Delta_t = (-1)^{t+1}D'(\rho_t)n_t/n.$$

It follows that

$$\begin{aligned} N/n^{\lambda-1} &= n^2\prod_{t=0}^{\lambda}(n_t/n) = \left(n^2\prod_{t=0}^{\lambda}F_t\right)\Big/\prod_{t=0}^{\lambda}(-1)^{t+1}D'(\rho_t) \\ &= \left(n^2\prod_{t=0}^{\lambda}F_t\right)\Big/\Delta^2 = \left(\prod_{t=0}^{\lambda}F_t\right)\Big/F_0^2\Delta_0^2. \end{aligned} \tag{2.11}$$

The proof of Theorem I depends now only on the proof of

LEMMA I. $\prod_{t=0}^{\lambda} F_t = KF_0^2F^2$, *where* $K = k_1k_2\cdots k_\lambda$, *and* F *is a factor of* Δ_0 *such that* $R = \Delta_0/F$ *is an integer.*

3. We shall now prove the theorem for the cases $\lambda = 2$, $\lambda = 3$. To facilitate the algebraic work we introduce the notation:

$$\theta_i = \rho_0 - \rho_i;\quad \phi_0 = 1,\quad \phi_1 = \sum\theta_i,\quad \phi_2 = \sum\theta_i\theta_j,\ \cdots,\ \phi_\lambda = \theta_1\theta_2\cdots\theta_\lambda. \tag{3.1}$$

$$\sigma_0 = 1 - 1/n,\quad \sigma_1 = s_1 - \rho_0,\quad \sigma_\mu = \sum_t\binom{\mu}{t}(-\rho_0)^t s_{\mu-t},\quad \mu > 0. \tag{3.2}$$

Since

$$F_0 = \begin{vmatrix} 1 & 1 & \cdots & 1 \\ \sigma_1 & -\theta_1 & \cdots & -\theta_\lambda \\ \sigma_2 & \theta_1^2 & \cdots & \theta_\lambda^2 \\ \cdot & \cdot & \cdots & \cdot \\ \sigma_\lambda & (-\theta_1)^\lambda & \cdots & (-\theta_\lambda)^\lambda \end{vmatrix} \div \begin{vmatrix} 1 & \cdots & 1 \\ -\theta_1 & \cdots & -\theta_\lambda \\ \theta_1^2 & \cdots & \theta_\lambda^2 \\ \cdot & \cdots & \cdot \\ (-\theta_1)^{\lambda-1} & \cdots & (-\theta_\lambda)^{\lambda-1} \end{vmatrix} \tag{3.3}$$

$$= (-1)^\lambda(\phi_\lambda + \sigma_1\phi_{\lambda-1} + \cdots + \sigma_\lambda\phi_0)$$

and

$$F_0 = (-1)^\lambda\phi_\lambda/n, \tag{3.4}$$

we have

$$\sum_{t=0}^{\lambda}\sigma_t\phi_{\lambda-t} = 0. \tag{3.5}$$

The formula for F_t, $t > 0$, becomes

$$F_t = \begin{vmatrix} 1 & 1 & \cdots & 1 & 1 & 1 & \cdots & 1 \\ 0 & -\theta_1 & \cdots & -\theta_{t-1} & \sigma_1 & -\theta_{t+1} & \cdots & -\theta_\lambda \\ 0 & \theta_1^2 & \cdots & \theta_{t-1}^2 & \sigma_2 & \theta_{t+1}^2 & \cdots & \theta_\lambda^2 \\ \cdot & \cdot & \cdots & \cdot & \cdot & \cdot & \cdots & \cdot \\ 0 & (-\theta_1)^\lambda & \cdots & (-\theta_{t-1})^\lambda & \sigma_\lambda & (-\theta_{t+1})^\lambda & \cdots & (-\theta_\lambda)^\lambda \end{vmatrix}$$

$$\div \begin{vmatrix} 1 & 1 & \cdots & 1 & 1 & \cdots & 1 \\ 0 & -\theta_1 & \cdots & -\theta_{t-1} & -\theta_{t+1} & \cdots & -\theta_\lambda \\ 0 & \theta_1^2 & \cdots & \theta_{t-1}^2 & \theta_{t+1}^2 & \cdots & \theta_\lambda^2 \\ \cdot & \cdot & \cdots & \cdot & \cdot & \cdots & \cdot \\ 0 & (-\theta_1)^{\lambda-1} & \cdots & (-\theta_{t-1})^{\lambda-1} & (-\theta_{t+1})^{\lambda-1} & \cdots & (-\theta_\lambda)^{\lambda-1} \end{vmatrix}$$

$$= (-1)^{\lambda-t}[\sigma_1\phi_{\lambda-1}^{(t)} + \sigma_2\phi_{\lambda-2}^{(t)} + \cdots + \sigma_\lambda],$$

where $\phi_\mu^{(t)}$ is obtained from ϕ_μ by setting $\theta_t = 0$, and satisfies the equation $\phi_\mu^{(t)} = \phi_\mu - \theta_t\phi_{\mu-1}^{(t)}$. In view of the identity (3.5), we have

$$F_t = (-1)^{\lambda-t}[-\sigma_0\phi_\lambda - \sigma_1(\phi_{\lambda-1} - \phi_{\lambda-1}^{(t)}) - \cdots - \sigma_{\lambda-1}(\phi_1 - \phi_1^{(t)})],$$

and hence

$$F_t = (-1)^{\lambda-t-1}\theta_t[\sigma_0\phi_{\lambda-1}^{(t)} + \sigma_1\phi_{\lambda-2}^{(t)} + \cdots + \sigma_{\lambda-1}\phi_0^{(t)}]. \tag{3.6}$$

Now consider the case $\lambda = 2$. We have

$$\sigma_0 = 1 - 1/n, \quad \sigma_1 = -(\alpha_1 k_1 + \alpha_2 k_2), \quad \sigma_2 = \sigma_1^2 + \alpha_1^2 k_1 + \alpha_2^2 k_2, \tag{3.2a}$$

$$F_0 = \theta_1\theta_2/n, \quad F_1 = \theta_1(\sigma_0\theta_2 + \sigma_1), \quad F_2 = -\theta_2(\sigma_0\theta_1 + \sigma_1), \tag{3.4a}$$

$$\begin{aligned} F_0F_1F_2 &= -F_0\theta_1\theta_2(\sigma_0^2\theta_1\theta_2 + \sigma_0\sigma_1(\theta_1 + \theta_2) + \sigma_1^2) \\ &= -F_0^2 n(-\sigma_0\sigma_2 + \sigma_1^2) = F_0^2[(n-1)\sigma_2 - n\sigma_1^2] \\ &= F_0^2[(k_1 + k_2)(\alpha_1^2 k_1 + \alpha_2^2 k_2) - (\alpha_1 k_1 + \alpha_2 k_2)^2], \end{aligned}$$

$$F_0F_1F_2 = F_0^2(k_1k_2)(\alpha_1 - \alpha_2)^2 = KF_0^2F^2, \tag{3.7a}$$

where

$$F = (\alpha_1 - \alpha_2). \tag{3.8a}$$

To prove the lemma in this case, we show that F divides Δ_0.

$$D(\rho) = \begin{vmatrix} \alpha_0 - \rho & \alpha_1 & \alpha_2 \\ \alpha_1 k_1 & C_{11} - \rho & C_{12} \\ \alpha_2 k_2 & C_{21} & C_{22} - \rho \end{vmatrix} = \begin{vmatrix} \alpha_0 - \rho & \alpha_1 & \alpha_2 \\ \alpha_1 k_1 & C_{11} - \rho & C_{12} \\ \rho_0 - \rho & \rho_0 - \rho & \rho_0 - \rho \end{vmatrix} \tag{3.9}$$

$$= (\rho_0 - \rho) \begin{vmatrix} \alpha_0 - \alpha_2 - \rho & \alpha_1 - \alpha_2 \\ \alpha_1 k_1 - C_{12} & C_{11} - C_{12} - \rho \end{vmatrix} = (\rho_0 - \rho)(\rho_1 - \rho)(\rho_2 - \rho).$$

$$\Delta^2/n^2F_0^2 = \Delta_0^2 = (\rho_1 - \rho_2)^2 = (\rho_1 + \rho_2)^2 - 4\rho_1\rho_2$$

$$= (\alpha_0 - \alpha_2 + C_{11} - C_{12})^2 - 4(\alpha_0 - \alpha_2)(C_{11} - C_{12}) + 4(\alpha_1 - \alpha_2)(\alpha_1 k_1 - C_{12})$$

$$= (-\alpha_2 - \alpha_1 c_{111} - \alpha_2 c_{211} + \alpha_1 c_{112} + \alpha_2 c_{212})^2 + 4(\alpha_1 - \alpha_2)(\alpha_1 k_1 - \alpha_1 c_{112} - \alpha_2 c_{212})$$

$$= (\alpha_1 - \alpha_2)^2 (c_{112} - c_{111})^2 + 4(\alpha_1 - \alpha_2)^2 c_{212}.$$

$$\Delta_0^2 = (\alpha_1 - \alpha_2)^2[(c_{112} - c_{111})^2 + 4c_{212}] = F^2R^2. \tag{3.10}$$

This last bracket will be a perfect square

$$R^2 = (c_{112} - c_{111})^2 + 4c_{212} \tag{3.11}$$

if $k_1 \neq k_2$, since ρ_1 and ρ_2 will then be rational. Lemma I follows from (3.7a), (3.8a), (3.10), and (3.11), so the theorem is proved in this case.

The algebraic work, even for the case $\lambda = 3$, is quite complicated. Using the same notation, we have

$$\begin{gathered} \sigma_0 = 1 - 1/n, \qquad \sigma_1 = -(\alpha_1 k_1 + \alpha_2 k_2 + \alpha_3 k_3), \\ \sigma_2 = \sigma_1^2 + \alpha_1^2 k_1 + \alpha_2^2 k_2 + \alpha_3^2 k_3, \qquad \sigma_0\phi_3 + \sigma_1\phi_2 + \sigma_2\phi_1 + \sigma_3 = 0. \end{gathered} \tag{3.2b}$$

(3.4b), (3.6b) $\quad F_0 = -\theta_1\theta_2\theta_3/n, \qquad F_1 = -\theta_1[\sigma_0\theta_2\theta_3 + \sigma_1(\theta_2 + \theta_3) + \sigma_2]$, etc.

$$\begin{aligned} F_1F_2F_3/F_0 = -n[&\sigma_0^3\phi_3^3 + 2\sigma_0^2\sigma_1\phi_2\phi_3 + \sigma_0^2\sigma_2\phi_1\phi_3 + \sigma_0\sigma_1^2(\phi_2^2 + \phi_1\phi_3) \\ &+ \sigma_0\sigma_1\sigma_2(3\phi_3 + \phi_1\phi_2) + \sigma_1^3(\phi_2\phi_1 - \phi_3) + \sigma_1^2\sigma_2(\phi_2 + \phi_1^2) + 2\sigma_1\sigma_2^2\phi_1 + \sigma_2^3]. \end{aligned} \tag{3.7b}$$

The left hand side may be simplified, and can be shown to equal KF^2, where

$$\begin{aligned} F = (\alpha_2 - \alpha_3)(\alpha_2 - \alpha_1)(C_{23} - C_{21}) + (\alpha_3 - \alpha_1)(\alpha_3 - \alpha_2)(C_{31} - C_{32}) \\ + (\alpha_1 - \alpha_2)(\alpha_1 - \alpha_3)(C_{12} - C_{13}) - (\alpha_1 - \alpha_2)(\alpha_2 - \alpha_3)(\alpha_3 - \alpha_1). \end{aligned} \tag{3.8b}$$

By (2.11), we have $N/n^{\lambda-1} = KF^2/\Delta_0^2$, or $Kn^{\lambda-1}/N = \Delta_0^2/F^2$. If the numbers k_i are distinct, the roots $\rho_1 \cdots \rho_\lambda$ must be rational, and so must Δ_0, the product

of their differences. Δ_0/F is therefore a rational fraction, and is independent of the parameters α_i. Assume it reduced to lowest terms. For $\alpha_2 = \alpha_3 = 0$, we see that the denominator must be a factor of $c_{112} - c_{113}$; similarly, it must be a factor of $c_{223} - c_{221}$, and of $c_{331} - c_{332}$, and hence a common factor of these three. But since we can choose α_1, α_2, α_3 so that F has no factor in common with these three, the denominator must be 1, and $\Delta_0/F = R$ is a rational integer.

4. We shall now consider an infinite family of simply transitive groups for which we can compute the constants n_i, and shall show that our formula in Theorem I holds also for larger values of λ than $\lambda = 3$. We define these groups as follows. Under the symmetric group of degree $\nu = \lambda + \mu$, the $n = \binom{\nu}{\lambda}$ combinations of the ν symbols taken λ at a time will be permuted among themselves by a transitive group G^*. We may denote these combinations by variables $x_{i_1 i_2 \cdots i_\lambda}$, defined to be independent of the order of the subscripts. The subgroup G_1 leaving fixed the variable $x_{12\cdots\lambda}$ permutes transitively among themselves those variables which have exactly $\lambda - t$ of the subscripts $1, 2, \cdots, \lambda$, and t of the remaining μ subscripts. Hence we have $r = \lambda + 1$ sets of transitivity T_t, $t = 0, 1, 2, \cdots, \lambda$, which contain respectively $k_t = \binom{\lambda}{t}\binom{\mu}{t}$ variables. The structure coefficient c_{abt} already defined in (2.2) for the ring of the matrices $V = \sum_{i=0}^{\lambda} \alpha_i V_i$ is the coefficient of $\alpha_a \alpha_b$ in the inner product of the row of V corresponding to the variable $x_{1,2,\cdots,\lambda}$ and the column corresponding to the variable $x_{t+1,t+2,\cdots,t+\lambda}$. It is given explicitly by the formula

$$c_{abt} = \sum_s \binom{t}{a-s}\binom{t}{b-s}\binom{\lambda-t}{s}\binom{\mu-t}{a+b-t-s}. \tag{4.1}$$

We verify that

$$c_{ab0} = \sum_s \binom{0}{a-s}\binom{0}{b-s}\binom{\lambda}{s}\binom{\mu}{a+b-s} = \binom{0}{b-a}\binom{\lambda}{a}\binom{\mu}{b} = k_a\delta_{ab}, \tag{4.1a}$$

$$c_{a0t} = \binom{t}{a}\binom{\mu-t}{a-t} = \delta_{at}. \tag{4.1b}$$

Let the characteristic equation $\det\left(\sum_a \alpha_a c_{abt} - \rho\delta_{bt}\right) = 0$ have the roots

$$\rho_i = \sum_{t=0}^{\lambda} \rho_{it}\alpha_t. \tag{4.2}$$

Then by solving for simple cases, and generalizing the result, we surmise that

$$\rho_{it} = \sum_{u=0}^{i} (-1)^u \binom{i}{u}\binom{\lambda-i}{t-u}\binom{\mu-i}{t-u}; \qquad \rho_{0t} = k_t; \qquad \rho_{i0} = 1. \tag{4.3}$$

To prove that this formula is correct, it is sufficient to show that

$$\sum_{t=0}^{\lambda} c_{abt}\rho_{it} = \rho_{ia}\rho_{ib}, \tag{4.4}$$

since if we multiply by the corresponding ρ_{it} each column of $\det\left(\sum_a \alpha_a c_{abt} - \rho\delta_{bt}\right)$ and add it to the first, this first column then has in its $(b+1)$-th row the quantity

$$\sum_{a,t} \alpha_a c_{abt}\rho_{it} - \rho\rho_{ib} = \sum_a \alpha_a \rho_{ia}\rho_{ib} - \rho\rho_{ib} = (\rho_i - \rho)\rho_{ib}.$$

Hence the determinant vanishes for $\rho = \rho_i$. Equation (4.4) follows from the rather complicated identity

LEMMA II. $$\sum_{s,t,u} \binom{t}{a-s}\binom{t}{b-s}\binom{\lambda-t}{s}\binom{\mu-t}{a+b-t-s}(-1)^u\binom{i}{u}\binom{\lambda-i}{t-u}\binom{\mu-i}{t-u} = \sum_p (-1)^p\binom{i}{p}\binom{\lambda-i}{a-p}\binom{\mu-i}{a-p}\sum_q (-1)^q\binom{i}{q}\binom{\lambda-i}{b-q}\binom{\mu-i}{b-q}$$

for all λ, μ, i, a, b; the summation being taken over all values of the arguments for which the binomial coefficients are different from zero.

In order not to interrupt the continuity, we postpone the proof of Lemma II until §5, and proceed to determine the constants n_i. The equation (2.8), for $h = 1$, is $\sum_{i=0}^{\lambda} n_i\rho_i = n\alpha_0$, and actually involves $\lambda + 1$ equations

$$\sum_{i=0}^{\lambda+1} n_i\rho_{it} = n\delta_{0t}, \qquad t = 0, 1, \cdots, \lambda, \tag{4.5}$$

which can be solved uniquely for the constants n_i. The determinant $\mathrm{P} = \det(\rho_{it})$ has the value

$$\mathrm{P} = (-1)^{\lambda+1}\prod_{j=0}^{\lambda}\binom{\nu-2j}{\lambda-j} \tag{4.6}$$

and

$$n_i = \binom{\nu}{i} - \binom{\nu}{i-1} = (\nu+1-2i)\nu!/(\nu+1-i)!\,i!. \tag{4.7}$$

We may now verify Theorem I for these groups. We have

$$N = \prod_{t=1}^{\lambda} n_t = \prod_{t=1}^{\lambda}\{(\nu+1-2t)\nu!/(\nu+1-t)!\,t!\}$$

$$n^{\lambda}K = \prod_{t=1}^{\lambda} nk_t = \prod_{t=1}^{\lambda}\{\nu!/t!\,(\lambda-t)!\,t!\,(\nu-\lambda-t)!\}$$

$$n^{\lambda}K/N = \prod_{t=1}^{\lambda}\{(\nu+1-t)!/(\lambda-t)!\,t!\,(\nu-\lambda-t)!\,(\nu+1-2t)\}$$

(4.8) $n^{\lambda-1}K/N = \prod_{t=1}^{\lambda} \{(\nu - t)!/(\lambda - t)!(t-1)!(\nu - \lambda - t)!(\nu + 1 - 2t)\} = R^2,$

where

(4.9) $$R = \left[\prod_{t=1}^{\lambda} \{1/(t-1)!\}\right] \prod_{s=1}^{[\lambda/2]} \{(\nu - 2s)!/(\nu - 1 - 2\lambda + 2s)!\}.$$

The quantity R is easily shown to be an integer, and the theorem is proved for this family of groups.

5. We close with a proof of Lemma II, which in itself involves interesting relations between binomial coefficients. Use is made of the following identities, which may be readily verified.

(5.α) $$\sum_\sigma (-1)^\sigma \binom{A}{B+\sigma}\binom{C+\sigma}{D} = (-1)^{A-B}\binom{C-B}{D-A}.$$

(5.β) $$\sum_\sigma \binom{A}{\sigma}\binom{\sigma}{B}\binom{B+C}{B+D-\sigma} = \binom{A+C}{D}\binom{A}{B}.$$

(5.γ) $$\sum_\sigma \binom{A}{B+\sigma}\binom{C}{D-\sigma} = \binom{A+C}{B+D}.$$

Proof of Lemma II.

(5.1) $$\sum_{s,t,u} \binom{t}{a-s}\binom{t}{b-s}\binom{\lambda-t}{s}\binom{\mu-t}{a+b-t-s}(-1)^u\binom{i}{u}\binom{\lambda-i}{t-u}\binom{\mu-i}{t-u},$$

(5.2) $$= \sum_{s,u,v} (-1)^u \binom{i}{u}\binom{u+v}{a-s}\binom{u+v}{b-s}\binom{\lambda-u-v}{s}\binom{\lambda-i}{v}$$
$$\binom{\mu-u-v}{a+b-s-u-v}\binom{\mu-i}{v}, \text{ where } v = t-u,$$

(5.3) $$= \sum_{\substack{p,q,r,s,\\u,v,x,y}} (-1)^u\binom{i}{u}\left\{\binom{u+v+s-x-q}{p-r}\binom{-v-s+x+q}{-v-s+x+r}\binom{v}{x-a+p}\right\}$$
$$\left\{\binom{u}{-v-s+x+q}\binom{v}{x-b+q}\right\}\left\{\binom{\lambda-i}{x}\binom{x}{v}\binom{i-u}{s+v-x}\right\}$$
$$\left\{\binom{\mu-i}{y}\binom{y}{v}\binom{i-u}{a+b-s-u-y}\right\}$$

by repeated application of (5.γ) and (5.β). Now if we rearrange the factorials in these binomial coefficients in a different way, and set

$$f(v; p, q, x, y) = \binom{\lambda-i}{x}\binom{x}{v}\binom{v}{x-a+p}\binom{\mu-i}{y}\binom{y}{v}\binom{v}{x-b+q},$$

we obtain for (5.3) the expression

$$\sum_{\substack{p,q,r,s,\\u,v,x,y}} (-1)^u \binom{i-p-q+r}{i-u-v-s+x} \binom{i-u}{i-a-b+s+y} \binom{r}{s+v-x} \binom{q}{r} \binom{i-q}{p-r} \binom{i}{q} f(v;\, p,\, q,\, x,\, y), \tag{5.4}$$

$$= \sum_{\substack{p,q,r,s,\\v,x,y}} (-1)^{p+q-r+s+v-x} \binom{s+v-x}{p+q-r-a-b+s+y} \binom{r}{s+v-x} \binom{q}{r} \binom{i-q}{p-r} \binom{i}{q} f(v;\, p,\, q,\, x,\, y), \tag{5.5}$$

$$= \sum_{p,q,r,v,x,y} (-1)^{p+q} \binom{0}{-p-q+a+b+v-x-y} \binom{q}{r} \binom{i-q}{p-r} \binom{i}{q} f(v;\, p,\, q,\, x,\, y), \tag{5.6}$$

$$= \sum_{p,q,x,y} (-1)^{p+q} \binom{i}{p} \binom{i}{q} f(p+q+x+y-a-b;\, p,\, q,\, x,\, y), \tag{5.7}$$

$$= \sum_{p,q,x,y} (-1)^{p+q} \binom{i}{p} \binom{i}{q} \binom{\lambda-i}{a-p} \binom{\lambda-i-a+p}{x-a+p} \binom{a-p}{a+b-p-q-x} \binom{\mu-i}{a-p} \binom{\mu-i-a+p}{y-a+p} \binom{a-p}{a+b-p-q-y}, \tag{5.8}$$

$$= \sum_{p,q} (-1)^{p+q} \binom{i}{p} \binom{i}{q} \binom{\lambda-i}{a-p} \binom{\lambda-i}{b-q} \binom{\mu-i}{a-p} \binom{\mu-i}{b-q}, \tag{5.9}$$

$$= \sum_{p} (-1)^{p} \binom{i}{p} \binom{\lambda-i}{a-p} \binom{\mu-i}{a-p} \sum_{q} (-1)^{q} \binom{i}{q} \binom{\lambda-i}{b-q} \binom{\mu-i}{b-q}, \tag{5.10}$$

which proves the lemma.

Brown University.

STRESSES IN MODERATELY THICK RECTANGULAR PLATES

R. ARCHIE HIGDON AND D. L. HOLL

1. Introduction

In 1922 G. D. Birkhoff[1] suggested a method for solving plate problems which involves the representation of the displacements by power series. C. A. Garabedian[2] and H. W. Sibert[3] used this idea in developing methods for solving problems in moderately thick circular plates. Garabedian[4] has also published some results for uniformly loaded rectangular plates.

The authors give a solution for the displacements in an elastic isotropic moderately thick rectangular plate under the action of any given load which can be expressed as a polynomial in x, y continuous over the entire plate and with prescribed boundary conditions at the edges. The method, similar to that used by Sibert[5] for circular plates, is based on the assumption that the components of displacement can be developed in positive integral powers of z. In this type of problem, the displacements must satisfy (a) the stress equations of equilibrium throughout the plate, (b) the surface traction conditions on the upper and lower faces, (c) the boundary conditions at the edges.

2. General theory

a. **Form of the displacements.** The displacements, u, v, w, are given by

$$(1) \qquad u = \sum_{n=0}^{\infty} U_n \frac{z^n}{n!}, \qquad v = \sum_{n=0}^{\infty} V_n \frac{z^n}{n!}, \qquad w = \sum_{n=0}^{\infty} W_n \frac{z^n}{n!},$$

where U_n, V_n, and W_n are continuous and continuously differentiable functions of x, y. The equations of equilibrium are (A. E. H. Love,[6] p. 134)

$$(2) \qquad (\lambda + \mu)\left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right)\Delta + \mu\nabla^2(u, v, w) = 0.$$

Received July 15, 1936; presented to the American Mathematical Society, April 11, 1936.

[1] G. D. Birkhoff, *Circular plates of variable thickness*, Phil. Mag., vol. 43 (1922), pp. 953–962.

[2] C. A. Garabedian, *Circular plates of constant or variable thickness*, Trans. Amer. Math. Soc., vol. 25 (1923), pp. 343–398.

[3] H. W. Sibert, *Moderately thick circular plates with plane faces*, Trans. Amer. Math. Soc., vol. 33 (1931), pp. 329–369.

[4] C. A. Garabedian, Comptes Rendus, Paris, vols. 178 (1924), 180 (1925), 181 (1925), 186 (1928), 195 (1932).

[5] Loc. cit.

[6] In this paper all references to Love are to the fourth edition of his *Mathematical Theory of Elasticity*, 1927.

When equations (1) are substituted in (2) and the coefficients of like powers of z are equated, there results[7]

$$(3.1)\qquad U_n = \frac{-1}{1-2\sigma}\left[(1-2\sigma)\nabla^2 U_{n-2} + \frac{\partial^2 U_{n-2}}{\partial x^2} + \frac{\partial^2 V_{n-2}}{\partial x \partial y} + \frac{\partial W_{n-1}}{\partial x}\right],$$

$$(3.3)^8\qquad W_n = \frac{-1}{2(1-\sigma)}\left[(1-2\sigma)\nabla^2 W_{n-2} + \frac{\partial U_{n-1}}{\partial x} + \frac{\partial V_{n-1}}{\partial y}\right].$$

By successive applications of the recurrence relations (3), U_n, V_n, and W_n are expressed in terms of U_0, V_0, W_0, U_1, V_1, and W_1. The results are

$$(4.1)\qquad U_{2k} = (-1)^k \nabla^{2k-2}\left[\nabla^2 U_0 + k\frac{\partial \overline{W}_1}{\partial x}\right],$$

$$(4.3)\qquad W_{2k+1} = (-1)^k \nabla^{2k}\left[(1-2\sigma-k)\overline{W}_1 - \left(\frac{\partial U_0}{\partial x} + \frac{\partial V_0}{\partial y}\right)\right],$$

$$(4.4)\qquad U_{2k+1} = (-1)^k \nabla^{2k-2}\left[\nabla^2 U_1 + k\frac{\partial \overline{W}_0}{\partial x}\right],$$

$$(4.6)\qquad W_{2k} = (-1)^k \nabla^{2k-2}\left[(k-2+2\sigma)\overline{W}_0 + \left(\frac{\partial U_1}{\partial x} + \frac{\partial V_1}{\partial y}\right)\right],$$

where

$$(1-2\sigma)\overline{W}_1 = \frac{\partial U_0}{\partial x} + \frac{\partial V_0}{\partial y} + W_1,$$

$$2(1-\sigma)\overline{W}_0 = \frac{\partial U_1}{\partial x} + \frac{\partial V_1}{\partial y} - \nabla^2 W_0 \qquad (k = 0, 1, 2, 3, \cdots).$$

When formulas (4) are substituted in (1) there results

$$(5.1)\qquad \begin{aligned} u = &\sum_{k=0}^{\infty} (-1)^k \nabla^{2k-2}\left[\nabla^2 U_0 + k\frac{\partial \overline{W}_1}{\partial x}\right]\frac{z^{2k}}{(2k)!} \\ &+ \sum_{k=0}^{\infty} (-1)^k \nabla^{2k-2}\left[\nabla^2 U_1 + k\frac{\partial \overline{W}_0}{\partial x}\right]\frac{z^{2k+1}}{(2k+1)!}, \end{aligned}$$

$$(5.3)\qquad \begin{aligned} w = &\sum_{k=0}^{\infty} (-1)^k \nabla^{2k-2}\left[(k-2+2\sigma)\overline{W}_0 + \left(\frac{\partial U_1}{\partial x} + \frac{\partial V_1}{\partial y}\right)\right]\frac{z^{2k}}{(2k)!} \\ &+ \sum_{k=0}^{\infty} (-1)^k \nabla^{2k}\left[(1-2\sigma-k)\overline{W}_1 - \left(\frac{\partial U_0}{\partial x} + \frac{\partial V_0}{\partial y}\right)\right]\frac{z^{2k+1}}{(2k+1)!}. \end{aligned}$$

These expressions for the displacements satisfy formally the stress equations of equilibrium throughout the body.

[7] Here and henceforward $\nabla^2 = \dfrac{\partial^2}{\partial x^2} + \dfrac{\partial^2}{\partial y^2}$.

[8] (3.2) is omitted because it can be obtained from (3.1) by interchanging x and y, U and V, etc. Throughout this paper an equation will be omitted when it can be obtained from the preceding equation in this manner.

b. **Surface traction equations.** A right-handed coördinate system with the faces of the plate $z = \pm h$, $x = 0$, $x = a$, $y = 0$, and $y = b$ will be used. The x, y, and z components of the surface tractions will be designated by L_1, J_1, and P_1 respectively on the upper face and by L_2, J_2, and P_2 respectively on the lower face. Using the notation of Love (p. 77), the surface traction conditions on the upper and lower faces may be written as

$$
\begin{aligned}
(X_z)_{z=h} &= L_1, & (X_z)_{z=-h} &= L_2, \\
(Y_z)_{z=h} &= J_1, & (Y_z)_{z=-h} &= J_2, \\
(Z_z)_{z=h} &= P_1, & (Z_z)_{z=-h} &= P_2 .
\end{aligned} \tag{6}
$$

The stresses in terms of the displacements are (Love, p. 101)

$$
\begin{aligned}
Z_z &= \frac{2\mu}{1-2\sigma}\left[(1-\sigma)\frac{\partial w}{\partial z} + \sigma\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}\right)\right], \\
X_z &= \mu\left(\frac{\partial u}{\partial z} + \frac{\partial w}{\partial x}\right), \qquad Y_z = \mu\left(\frac{\partial v}{\partial z} + \frac{\partial w}{\partial y}\right).
\end{aligned} \tag{7}
$$

Substitute equations (5) in (7) and the results in (6). Then take the sum and difference of the two resulting values of Z_z, X_z, and Y_z. The final equations are

$$
\begin{aligned}
\sum_{k=0}^{\infty} (-1)^k \nabla^{2k-2}\Bigg[\nabla^2 U_1 &+ \frac{k}{1-\sigma}\left(\frac{\partial^2 U_1}{\partial x^2} + \frac{\partial^2 V_1}{\partial x \partial y}\right) \\
&- \frac{k-1+\sigma}{1-\sigma}\nabla^2\left(\frac{\partial W_0}{\partial x}\right)\Bigg]\frac{h^{2k}}{(2k)!} = \frac{L_1 + L_2}{2\mu},
\end{aligned} \tag{8}
$$

$$
\begin{aligned}
\sum_{k=0}^{\infty} (-1)^k \nabla^{2k}\left[\frac{k-1+\sigma}{2(1-\sigma)}\nabla^2 W_0 - \frac{k+1-\sigma}{2(1-\sigma)}\left(\frac{\partial U_1}{\partial x} + \frac{\partial V_1}{\partial y}\right)\right]\frac{h^{2k}}{(2k+1)!} \\
= \frac{P_1 - P_2}{4\mu h},
\end{aligned} \tag{10}
$$

$$
\begin{aligned}
\sum_{k=0}^{\infty} (-1)^{k+1} \nabla^{2k}\Bigg[\nabla^2 U_0 &+ \frac{2k+1}{1-2\sigma}\left(\frac{\partial^2 U_0}{\partial x^2} + \frac{\partial^2 V_0}{\partial x \partial y}\right) \\
&+ \frac{2(k+\sigma)}{1-2\sigma}\frac{\partial W_1}{\partial x}\Bigg]\frac{h^{2k}}{(2k+1)!} = \frac{L_1 - L_2}{2\mu h},
\end{aligned} \tag{11}
$$

$$
\sum_{k=0}^{\infty} (-1)^{k+1} \nabla^{2k}\left[\frac{k-\sigma}{1-2\sigma}\left(\frac{\partial U_0}{\partial x} + \frac{\partial V_0}{\partial y}\right) + \frac{k-1+\sigma}{1-2\sigma} W_1\right]\frac{h^{2k}}{(2k)!} = \frac{P_1 + P_2}{4\mu}. \tag{13}
$$

Equations (8), (9) and (10) involve only W_0, U_1 and V_1 and equations (11), (12) and (13) involve only U_0, V_0 and W_1. These two simultaneous systems of equations can be solved by an indirect process due to Sibert.[9] This process

[9] Loc. cit.

requires that U_0, U_1, V_0, V_1, W_0 and W_1 be expressed as infinite sequences of terms of ascending order of magnitude. Let s represent a first degree function of x, y. Order of magnitude may then be defined as follows: If r and t are two functions of s which contain the same number of terms, t is defined to be of the n-th order of magnitude as compared to r if each term of t is proportional to $(h/s)^n$ times the corresponding term in r. It then follows that $h^{2n}\nabla^{2n}r$ is of the $2n$-th order of magnitude as compared to r. It is necessary to assume that U_0, V_0, W_0, U_1, V_1, and W_1 are expressions in x, y which involve h in such a manner that their terms can be grouped and arranged in ascending order of magnitude.

Since equations (8) to (13) inclusive have been arranged so that only even powers of h occur in their left members, it is only necessary to provide for even orders of magnitude. Therefore

$$U_0 = \sum_{n=0}^{\infty} U_{2n,0}, \qquad U_1 = \sum_{n=0}^{\infty} U_{2n,1}, \tag{14.1}$$

$$W_0 = \sum_{n=0}^{\infty} W_{2n,0}, \qquad W_1 = \sum_{n=0}^{\infty} W_{2n,1}, \tag{14.3}$$

where $W_{2n,0}$, $U_{2n,0}$, $V_{2n,0}$, $W_{2n,1}$, $U_{2n,1}$, and $V_{2n,1}$ are of the $2n$-th order of magnitude as compared to W_{00}, U_{00}, V_{00}, W_{01}, U_{01}, and V_{01} respectively. It is assumed that $W_{00}, \cdots, V_{01}$, being the terms of lowest order of magnitude, do not vanish identically unless $W_0, \cdots, V_1$ respectively are identically zero.

For simplicity the problem will now be restricted to the case of a normal surface load only. The solution for the case of a shearing load is very similar to this case. By superposing these solutions, the results for more complicated problems can be obtained.

c. **General solutions for W_0, U_1, and V_1.** In order to solve equations (8), (9), and (10) simultaneously it is necessary to write each one as an infinite system of equations by equating terms of the same order of magnitude. The right members of equations (8) and (9) are now zero, but the order of magnitude of the right member of (10) must be determined. Assume it to be of the same order of magnitude as $\nabla^2 W_{00}$. Let $P_1 - P_2 = -p_1$, where p_1 is a function of x, y. Then the equations of lowest order of magnitude in (8), (9), and (10) may be written

$$U_{01} + \frac{\partial W_{00}}{\partial x} = 0, \tag{8.0}$$

$$\nabla^2 W_{00} + \frac{\partial U_{01}}{\partial x} + \frac{\partial V_{01}}{\partial y} = \frac{p_1}{2\mu h}. \tag{10.0}$$

However $\dfrac{\partial(8.0)}{\partial x} + \dfrac{\partial(9.0)}{\partial y}$ gives $\nabla^2 W_{00} + \dfrac{\partial U_{01}}{\partial x} + \dfrac{\partial V_{01}}{\partial y} = 0$. This result is inconsistent with (10.0). Therefore $p_1/(2\mu h)$ must be of the fourth or higher order

of magnitude as compared to W_{00}. It will be assumed now and proved later that $p_1/(2\mu h)$ is of the fourth order of magnitude as compared with W_{00}. Equations (8), (9), and (10) may now be written as infinite systems as follows:

$$U_{01} + \frac{\partial W_{00}}{\partial x} = 0, \tag{8.0}$$

$$\begin{aligned} U_{2n,1} + \frac{\partial W_{2n,0}}{\partial x} + \sum_{k=1}^{n} (-1)^k \nabla^{2k-2}\Bigg[\nabla^2 U_{2n-2k,1} + \frac{k}{1-\sigma}\Bigg(\frac{\partial^2 U_{2n-2k,1}}{\partial x^2} \\ + \frac{\partial^2 V_{2n-2k,1}}{\partial x \partial y}\Bigg) - \frac{k-1+\sigma}{1-\sigma}\nabla^2\Bigg(\frac{\partial W_{2n-2k,0}}{\partial x}\Bigg)\Bigg]\frac{h^{2k}}{(2k)!} = 0 \quad (n = 1, 2, 3, \cdots). \end{aligned} \tag{8.n}$$

Since $\dfrac{\partial(8.0)}{\partial x} + \dfrac{\partial(9.0)}{\partial y} = (10.0)$, it is necessary to form another system of equations by subtracting $\dfrac{\partial(8.n)}{\partial x} + \dfrac{\partial(9.n)}{\partial y}$ from $(10.n)$. The resulting system is[10]

$$0 = 0, \tag{15.0}$$

$$\nabla^2\left[(2-\sigma)\left(\frac{\partial U_{01}}{\partial x} + \frac{\partial V_{01}}{\partial y}\right) - \sigma\nabla^2 W_{00}\right] = \frac{2p_1}{D}, \tag{15.1}$$

$$\begin{aligned} \sum_{k=1}^{n} (-1)^{k+1}\nabla^{2k}\Bigg[(k+1-\sigma)\Bigg(\frac{\partial U_{2n-2k,1}}{\partial x} + \frac{\partial V_{2n-2k,1}}{\partial y}\Bigg) \\ - (k-1+\sigma)\nabla^2 W_{2n-2k,0}\Bigg]\frac{6kh^{2k-2}}{(2k+1)!} = 0 \quad (n = 2, 3, 4, \cdots). \end{aligned} \tag{15.n}$$

Systems (8), (9), and (15) can now be solved simultaneously for $W_{2n,0}$, $U_{2n,1}$ and $V_{2n,1}$. Equations (15.1) and the Laplacian of $\dfrac{\partial(8.0)}{\partial x} + \dfrac{\partial(9.0)}{\partial y}$ become

$$\nabla^2\left[(2-\sigma)\left(\frac{\partial U_{01}}{\partial x} + \frac{\partial V_{01}}{\partial y}\right) - \sigma\nabla^2 W_{00}\right] = \frac{2b_0 p_1}{D},$$

$$\nabla^2\left[\left(\frac{\partial U_{01}}{\partial x} + \frac{\partial V_{01}}{\partial y}\right) + \nabla^2 W_{00}\right] = \frac{2d_0 p_1}{D},$$

respectively, where $b_0 = 1$ and $d_0 = 0$. The simultaneous solution is

$$\nabla^4 W_{00} = \frac{p_1}{D}[(2-\sigma)d_0 - b_0], \qquad \nabla^2\left(\frac{\partial U_{01}}{\partial x} + \frac{\partial V_{01}}{\partial y}\right) = \frac{p_1}{D}[\sigma d_0 + b_0].$$

[10] $D = \dfrac{4\mu h^3}{3(1-\sigma)} = \dfrac{2Eh^3}{3(1-\sigma^2)}$.

Equations (15.2) and the Laplacian of $\dfrac{\partial(8.1)}{\partial x} + \dfrac{\partial(9.1)}{\partial y}$ become

$$\nabla^2\left[(2-\sigma)\left(\frac{\partial U_{21}}{\partial x} + \frac{\partial V_{21}}{\partial y}\right) - \sigma\nabla^2 W_{20}\right] = \frac{2h^2 b_1 \nabla^2 p_1}{D},$$

$$\nabla^2\left[\left(\frac{\partial U_{21}}{\partial x} + \frac{\partial V_{21}}{\partial y}\right) + \nabla^2 W_{20}\right] = \frac{2h^2 d_1 \nabla^2 p_1}{D},$$

respectively, where $b_1 = \dfrac{1}{5}$ and $d_1 = \dfrac{1}{2(1-\sigma)}$. The simultaneous solution is

$$\nabla^4 W_{20} = \frac{h^2}{D}\nabla^2 p_1[(2-\sigma)d_1 - b_1], \qquad \nabla^2\left(\frac{\partial U_{21}}{\partial x} + \frac{\partial V_{21}}{\partial y}\right) = \frac{h^2}{D}\nabla^2 p_1[\sigma d_1 + b_1].$$

Continuing in this manner, by solving $(15.n+1)$ with the Laplacian of $\dfrac{\partial(8.n)}{\partial x} + \dfrac{\partial(9.n)}{\partial y}$, one obtains the following general solutions:

$$\nabla^4 W_{2n,0} = \frac{h^{2n}}{D}\nabla^{2n} p_1[(2-\sigma)d_n - b_n], \tag{16}$$

$$\nabla^2\left(\frac{\partial U_{2n,1}}{\partial x} + \frac{\partial V_{2n,1}}{\partial y}\right) = \frac{h^{2n}}{D}\nabla^{2n} p_1[\sigma d_n + b_n], \tag{17}$$

where

$$b_n = \sum_{i=0}^{n-1}(-1)^i \frac{6(i+2)\{(i+2)b_{n-1-i} - (i+1)(1-\sigma)d_{n-1-i}\}}{(2i+5)!}, \tag{18}$$

$$d_n = \sum_{i=0}^{n-1}(-1)^i \frac{(i+1)b_{n-1-i} - i(1-\sigma)d_{n-1-i}}{(1-\sigma)(2i+2)!} \qquad (n = 1, 2, 3, \cdots). \tag{19}$$

Sibert has given upper bounds of the sequences b_n and d_n as

$$|b_n| < \frac{1}{3^n}, \qquad |d_n| < \frac{1}{3^{n-1}4(1-\sigma)} \qquad (n = 2, 3, \cdots).$$

Since (16.0) is $\nabla^4 W_{00} = \dfrac{-p_1}{D}$, it is the differential equation defining the vertical displacement in thin plate theory (Love, p. 488), and W_{00} is the vertical displacement of the corresponding thin plate. Consequently the vertical displacement W_0 of the middle surface of a moderately thick plate is made up of the corresponding thin plate displacement W_{00} plus corrections W_{20}, W_{40}, etc.

By combining equations (16), (17), and (8), $U_{2n,1}$ is obtained in terms of $W_{2n,0}$. Similarly (16), (17), and (9) yield $V_{2n,1}$. The results are

$$U_{01} + \frac{\partial W_{00}}{\partial x} = 0, \tag{20.0}$$

$$\begin{aligned} &U_{2n,1} + \frac{\partial W_{2n,0}}{\partial x} + \frac{h^2}{1-\sigma}\nabla^2\left(\frac{\partial W_{2n-2,0}}{\partial x}\right) \\ &\quad + h^{2n}\nabla^{2n}\left(\frac{\partial W_{00}}{\partial x}\right)\left[2d_n + \frac{(2-\sigma)d_{n-1} - b_{n-1}}{1-\sigma}\right] = 0 \quad (n = 1, 2, 3, \cdots). \end{aligned} \tag{20.n}$$

If $W_{2n,0}$ is eliminated from equations (20) and (21) the relation $\frac{\partial U_{2n,1}}{\partial y} = \frac{\partial V_{2n,1}}{\partial x}$ is obtained. This relation combined with (17) gives

$$\nabla^4 U_{2n,1} = \frac{h^{2n}}{D}[\sigma d_n + b_n]\nabla^{2n}\left(\frac{\partial p_1}{\partial x}\right). \tag{22}$$

It is necessary to complete the proof that $p_1/(2\mu h)$ is of the fourth order of magnitude as compared to W_{00}. It has already been shown to be either of the fourth or of a higher order of magnitude. Assume that its order of magnitude as compared to W_{00} is greater than the fourth. Then equations (15.1) and the Laplacian of $\frac{\partial(8.0)}{\partial x} + \frac{\partial(9.0)}{\partial y}$ become

$$\nabla^2\left[(2-\sigma)\left(\frac{\partial U_{01}}{\partial x} + \frac{\partial V_{01}}{\partial y}\right) - \sigma\nabla^2 W_{00}\right] = 0,$$

$$\nabla^2\left[\left(\frac{\partial U_{01}}{\partial x} + \frac{\partial V_{01}}{\partial y}\right) + \nabla^2 W_{00}\right] = 0,$$

respectively. The simultaneous solution is

$$\nabla^4 W_{00} = \nabla^2\left(\frac{\partial U_{01}}{\partial x} + \frac{\partial V_{01}}{\partial y}\right) = 0.$$

Since the trivial solution $\nabla^4 W_{2n,0} = 0$ does not depend upon the load, p_1 must occur on the right side of some one of equations (16). But since (16.0) is the equation of lowest order of magnitude in the system (16), its right member cannot be zero unless the right members of all equations of the system are zero. From this contradiction it follows that the term $p_1/(2\mu h)$ must be of the fourth order of magnitude as compared to W_{00}.

Finally, equations (16), (17), (22), and (23) substituted in (14) give the following general solutions for W_0, U_1, and V_1:

$$\nabla^4 W_0 = \sum_{n=0}^{\infty} \frac{h^{2n}}{D}\nabla^{2n} p_1[(2-\sigma)d_n - b_n], \tag{24}$$

$$\nabla^4 U_1 = \sum_{n=0}^{\infty} \frac{h^{2n}}{D}\nabla^{2n}\left(\frac{\partial p_1}{\partial x}\right)[\sigma d_n + b_n], \tag{25}$$

$$\nabla^2\left(\frac{\partial U_1}{\partial x} + \frac{\partial V_1}{\partial y}\right) = \sum_{n=0}^{\infty} \frac{h^{2n}}{D}\nabla^{2n} p_1[\sigma d_n + b_n]. \tag{27}$$

In 1899 J. H. Michell[11] published the differential equation defining W_0 in the form

$$\nabla^4 W_0 = -\frac{(1-\sigma^2)}{E}\left[\left(\nabla^2 + \frac{\partial^2}{\partial z^2}\right)\frac{\partial Z_z}{\partial z}\right]_{z=0} - \frac{1+\sigma}{E}\nabla^2\left(\frac{\partial Z_z}{\partial z}\right)_{z=0}.$$

It is easily shown that this solution becomes identical with equation (24) when the stress Z_z is expressed in terms of the load.

d. **General solutions for W_1, U_0, and V_0.** Let $P_1 + P_2 = -p_2$, where p_2 is a function of x, y. It can be proved that the right member of (13) is of the same order of magnitude as W_{01}. Then equations (11), (12), and (13) can be written as infinite systems of equations. These systems can be solved in essentially the same manner as that used to obtain the simultaneous solution of systems (8), (9) and (15). The general solutions for U_0, V_0 and W_1 are

$$\nabla^2 W_1 = -\sum_{n=0}^{\infty}\frac{h^{2n}}{4\mu}\nabla^{2n+2}p_2[(1-\sigma)c_n + \sigma a_n], \tag{28}$$

$$\nabla^4 U_0 = \sum_{n=0}^{\infty}\frac{h^{2n}}{4\mu}\nabla^{2n+2}\left(\frac{\partial p_2}{\partial x}\right)[\sigma c_n + (1-\sigma)a_n], \tag{29}$$

$$\nabla^2\left(\frac{\partial U_0}{\partial x} + \frac{\partial V_0}{\partial y}\right) = \sum_{n=0}^{\infty}\frac{h^{2n}}{4\mu}\nabla^{2n+2}p_2[\sigma c_n + (1-\sigma)a_n], \tag{31}$$

where $a_0 = 0$, $c_0 = 1$,

$$\left.\begin{aligned} a_n &= \sum_{i=0}^{n-1}(-1)^i\frac{(i+2)a_{n-1-i} - (i+1)c_{n-1-i}}{(2i+3)!} &(32)\\ c_n &= \sum_{i=0}^{n-1}(-1)^i\frac{(i+1)a_{n-1-i} - ic_{n-1-i}}{(2i+2)!} &(33)\end{aligned}\right\}(n = 1, 2, \cdots).$$

Sibert has given as upper bounds for these sequences

$$|a_n| < \frac{1}{2^n}, \qquad |c_n| < \frac{5}{9\cdot 2^n} \qquad (n = 1, 2, 3\cdots).$$

The displacements u, v, w are given by relations (5) when the six functions U_0, V_0, W_0, U_1, V_1, and W_1 are known. Therefore one can say that the differential equations (24) to (31) inclusive define the displacements. Furthermore, the displacements defined by these differential equations satisfy the equilibrium equations and the surface traction conditions for any normal load which can be expressed as a polynomial in x, y continuous over the entire plate. It remains to solve these differential equations subject to particular sets of edge conditions.

[11] J. H. Michell, *On the direct determination of stress in an elastic solid with application to the theory of plates*, Proc. Lond. Math. Soc., vol. 31 (1899), pp. 100–124.

3. **Normal load,** $p_1(x, y) = p_2(x, y) = P\left(\lambda + \frac{\alpha x}{a} + \frac{\beta y}{b}\right)$

a. **Preliminary relations and remarks.** The displacements will now be found for a normal load $P\left(\lambda + \frac{\alpha x}{a} + \frac{\beta y}{b}\right)$ on the top surface of the plate where a and b are the horizontal dimensions of the plate, λ, α and β are arbitrary constants, and P is a uniform load per unit of area. In this case equations (24)–(31) become

$$\nabla^4 W_0 = \frac{-P}{D}\left(\lambda + \frac{\alpha x}{a} + \frac{\beta y}{b}\right), \tag{34}$$

$$\nabla^4 U_1 = \frac{P\alpha}{aD}, \tag{35}$$

$$\nabla^2\left(\frac{\partial U_1}{\partial x} + \frac{\partial V_1}{\partial y}\right) = \frac{P}{D}\left(\lambda + \frac{\alpha x}{a} + \frac{\beta y}{b}\right), \tag{37}$$

$$\nabla^2 W_1 = 0, \tag{38}$$

$$\nabla^4 U_0 = 0, \tag{39}$$

$$\nabla^2\left(\frac{\partial U_0}{\partial x} + \frac{\partial V_0}{\partial y}\right) = 0. \tag{41}$$

The load is a linear function of x, y and it is of the fourth order of magnitude as compared with W_{00}. Hence all terms whose orders of magnitude are $\geqq 6$ vanish. Therefore all infinite systems and sequences of the previous section become finite. The following results are readily deduced.

$$U_1 = -\frac{\partial W_0}{\partial x} - \frac{h^2}{1-\sigma}\nabla^2\left(\frac{\partial W_0}{\partial x}\right) - \frac{h^4(5-2\sigma)}{6(1-\sigma)^2}\nabla^4\left(\frac{\partial W_0}{\partial x}\right), \tag{42}$$

$$W_1 = \frac{-\sigma}{1-\sigma}\left(\frac{\partial U_0}{\partial x} + \frac{\partial V_0}{\partial y}\right) - \frac{P(1-2\sigma)}{4\mu(1-\sigma)}\left(\lambda + \frac{\alpha x}{a} + \frac{\beta y}{b}\right), \tag{44}$$

$$\begin{aligned} u = U_0 + U_1 z - \left[\nabla^2 U_0 + \frac{1}{1-2\sigma}\frac{\partial}{\partial x}\left(\frac{\partial U_0}{\partial x} + \frac{\partial V_0}{\partial y} + W_1\right)\right]\frac{z^2}{2!} \\ + \frac{2-\sigma}{1-\sigma}\nabla^2\left(\frac{\partial W_0}{\partial x}\right)\frac{z^3}{3!} + \frac{(3-2\sigma)h^2}{2(1-\sigma)^2}\nabla^4\left(\frac{\partial W_0}{\partial x}\right)\frac{z^3}{3!} - \frac{3-\sigma}{1-\sigma}\nabla^4\left(\frac{\partial W_0}{\partial x}\right)\frac{z^5}{5!}, \end{aligned} \tag{45}$$

$$w = W_0 + W_1 z + \frac{\sigma z^2}{2(1-\sigma)}\nabla^2 W_0 + \frac{h^2 z^2}{4(1-\sigma)^2}\nabla^4 W_0 - \frac{1+\sigma}{1-\sigma}\nabla^4 W_0 \frac{z^4}{4!}. \tag{47}$$

The problems of plane and generalized plane stress are very easily solved by use of equations (34) to (47) inclusive. For these problems $P = 0$, and since W_{01}, $\partial U_0/\partial x$ and $\partial V_0/\partial y$ are of the same order of magnitude as P, they may be chosen equal to zero. With these simplifications, the values of the stresses obtained from equations (45), (46) and (47) agree with those given by Love (pp. 467–471).

The problem of this investigation is restricted to moderately thick plates because the tractions applied to the edges are represented by their force- and couple-resultants taken along a vertical element of an edge. The boundary conditions at an edge will be defined in terms of the components of these resultants; namely, T, S, N, G, and H (Love, p. 455). At a pinned edge T, G, and W_0 vanish. At a free edge T, G, $N - \partial H/\partial s$ vanish, s denoting the direction along the edge. At a clamped edge U_0, V_0, W_0, and $\partial W_0/\partial n$ vanish, n denoting the direction of the normal to the edge line. These components must be expressed in terms of the displacements before they can be used here. The results are

$$(48)\quad \begin{aligned} T_1 = \frac{4\mu h}{1-\sigma}\Bigg[\frac{\partial U_0}{\partial x} + \sigma\frac{\partial V_0}{\partial y} - \frac{\sigma P}{4\mu}\left(\lambda + \frac{\alpha x}{a} + \frac{\beta y}{b}\right) \\ \frac{-h^2}{6}\left\{(1-\sigma)\nabla^2\left(\frac{\partial U_0}{\partial x}\right) + \frac{\partial^2}{\partial x^2}\left(\frac{\partial U_0}{\partial x} + \frac{\partial V_0}{\partial y}\right)\right\}\Bigg],\end{aligned}$$

$$(50)\quad G_1 = \frac{-4\mu h^3}{3(1-\sigma)}\left[\frac{\partial^2 W_0}{\partial x^2} + \sigma\frac{\partial^2 W_0}{\partial y^2} + \frac{8-3\sigma}{10(1-\sigma)}h^2\nabla^4 W_0 - \frac{8+\sigma}{10}h^2\nabla^2\left(\frac{\partial^2 W_0}{\partial y^2}\right)\right],$$

$$(52)\quad \begin{aligned} N_2 - \frac{\partial H_2}{\partial x} = \frac{-4\mu h^3}{3(1-\sigma)}\Bigg[(2-\sigma)\frac{\partial^3 W_0}{\partial x^2\,\partial y} + \frac{\partial^3 W_0}{\partial y^3} \\ + \frac{8+\sigma}{10}h^2\nabla^2\left(\frac{\partial^3 W_0}{\partial x^2\,\partial y}\right) + \frac{(8-3\sigma)h^2}{10(1-\sigma)}\nabla^4\left(\frac{\partial W_0}{\partial y}\right)\Bigg].\end{aligned}$$

Further, each equation which results from equating quantities (48) to (52) to zero must be written as a system of equations by equating terms of like orders of magnitude. The equation of lowest order of magnitude in each case will be referred to as (48.0), (49.0), $\cdots$, (52.0); the next order will be referred to as (48.1), (49.1), $\cdots$, (52.1), etc.

The method of solution used here does not give W_0 directly. First W_{00} is obtained, then W_{20}, etc., and finally W_0 is obtained from equation (14). It has already been pointed out that W_{00} is the vertical displacement for the corresponding thin plate under the same normal surface load. By observing that conditions (50.0), (51.0) and (52.0) are precisely the corresponding edge conditions from thin plate theory, one may employ the W_{00} solution of the thin plate problem in case it is available. The thin plate solution for the particular load of this problem has not been recorded and it will now be obtained for three sets of edge conditions.

b. **Solution for W_{00}.** The following Fourier expansions which are valid in the interval $0 < x < a$ will be needed.

$$(53)\quad \frac{\lambda}{2}(ax - x^2) = \frac{4\lambda}{a}\sum_{n=1,3,5,\cdots}\frac{\sin\theta x}{\theta^3} = \frac{2}{a}\sum_{n=1}^{\infty} i\,\frac{\sin\theta x}{\theta^3},$$

$$(54)\quad \frac{\alpha}{6a}(a^2x - x^3) = \frac{2\alpha}{a}\sum_{n=1}^{\infty}(-1)^{n+1}\frac{\sin\theta x}{\theta^3} = \frac{2}{a}\sum_{n=1}^{\infty} j\,\frac{\sin\theta x}{\theta^3},$$

$$\lambda + \frac{\alpha x}{a} = \frac{2}{a} \sum_{n=1}^{\infty} (i + j) \frac{\sin \theta x}{\theta}, \tag{55}$$

$$\frac{\lambda a^4}{24} \left(\frac{x^4}{a^4} - \frac{2x^3}{a^3} + \frac{x}{a} \right) = \frac{2}{a} \sum_{n=1}^{\infty} i \frac{\sin \theta x}{\theta^5}, \tag{56}$$

$$\frac{\alpha a^4}{24} \left(\frac{x^5}{5a^5} - \frac{2x^3}{3a^3} + \frac{7x}{15a} \right) = \frac{2}{a} \sum_{n=1}^{\infty} j \frac{\sin \theta x}{\theta^5}, \tag{57}$$

where

$$\theta = \frac{n\pi}{a}, \qquad i = \lambda[1 - (-1)^n], \qquad j = \alpha(-1)^{n+1}, \qquad m = \beta[1 - (-1)^n].$$

In the problems to be considered, the edges $x = 0, a$ will always be pinned. A solution satisfying edge conditions $W_{00} = 0$ and (50.0) at $x = 0, a$ and satisfying the differential equation (16.0) or (34.0) is given by

$$\begin{aligned} W_{00} = \frac{-Pa^4}{24D} \left[\left(\lambda + \frac{\beta y}{b} \right) \left(\frac{x^4}{a^4} - \frac{2x^3}{a^3} + \frac{x}{a} \right) + \alpha \left(\frac{x^5}{5a^5} - \frac{2x^3}{3a^3} + \frac{7x}{15a} \right) \right] \\ \frac{-P}{aD} \sum_n \frac{\sin \theta x}{\theta^5} [A_n \operatorname{sh} \theta y + B_n \operatorname{ch} \theta y + C_n \theta y \operatorname{sh} \theta y + D_n \theta y \operatorname{ch} \theta y], \end{aligned} \tag{58}$$

where the constants A_n, B_n, C_n and D_n are to be determined by the conditions at $y = 0, b$.

CASE I. *The edges $y = 0, b$ are pinned.* The boundary conditions are $W_{00} = 0$ and (51.0). When (58) is substituted in these edge conditions and use is made of the appropriate Fourier expansions, the following system of equations is obtained.

$$\begin{aligned} &B_n = -2(i + j), \\ &A_n \operatorname{sh} \phi + B_n \operatorname{ch} \phi + C_n \phi \operatorname{sh} \phi + D_n \phi \operatorname{ch} \phi = -2(i + j + m), \\ &B_n(1 - \sigma) + 2C_n = 2\sigma(i + j), \\ &A_n(1 - \sigma) \operatorname{sh} \phi + B_n(1 - \sigma) \operatorname{ch} \phi + C_n[\phi(1 - \sigma) \operatorname{sh} \phi + 2 \operatorname{ch} \phi] \\ &\qquad + D_n[\phi(1 - \sigma) \operatorname{ch} \phi + 2 \operatorname{sh} \phi] = 2\sigma(i + j + m), \end{aligned}$$

where $\phi = \theta b = n\pi b/a$. The simultaneous solution of this system of equations is

$$\begin{aligned} A_n &= -\frac{(i + j)(1 - \operatorname{ch} \phi)(2 \operatorname{sh} \phi - \phi) + m(2 \operatorname{sh} \phi + \phi \operatorname{ch} \phi)}{\operatorname{sh}^2 \phi}, \\ B_n &= -2(i + j), \\ C_n &= (i + j), \\ D_n &= \frac{(i + j)(1 - \operatorname{ch} \phi) + m}{\operatorname{sh} \phi}. \end{aligned}$$

These values substituted in (58) constitute the solution for W_{00} for $y = 0, b$ pinned. When $\lambda = 1$, this solution reduces to that published by S. Iguchi.[12]

CASE II. *The edges $y = 0, b$ are free.* The boundary conditions are (51.0) and (52.0). The substitution of equations (53), (54) and (58) in these edge conditions yields

$$B_n(1 - \sigma) + 2C_n = 2\sigma(i + j),$$

$$A_n(1 - \sigma)\operatorname{sh}\phi + B_n(1 - \sigma)\operatorname{ch}\phi + C_n[2\operatorname{ch}\phi + (1 - \sigma)\phi\operatorname{sh}\phi] + D_n[2\operatorname{sh}\phi + (1 - \sigma)\phi\operatorname{ch}\phi] = 2\sigma(i + j + m),$$

$$A_n\phi(1 - \sigma) - D_n\phi(1 + \sigma) = -2m(2 - \sigma),$$

$$A_n\phi(1 - \sigma)\operatorname{ch}\phi + B_n\phi(1 - \sigma)\operatorname{sh}\phi + C_n\phi[(1 - \sigma)\phi\operatorname{ch}\phi - (1 + \sigma)\operatorname{sh}\phi] + D_n\phi[(1 - \sigma)\phi\operatorname{sh}\phi - (1 + \sigma)\operatorname{ch}\phi] = -2m(2 - \sigma).$$

The simultaneous solution of this system of equations is

$$A_n = \frac{2\sigma(1 + \sigma)(i + j)(1 - \operatorname{ch}\phi)}{(1 - \sigma)(R - S)} + \frac{2m[\sigma(1 + \sigma)\phi(R - S\operatorname{ch}\phi) - (2 - \sigma)(R - S)(2\operatorname{sh}\phi + S)]}{S(R^2 - S^2)},$$

$$B_n = \frac{2\sigma(i + j)[(1 + \sigma)\operatorname{sh}\phi - S]}{(1 - \sigma)(R - S)} - \frac{4m[\sigma\phi S\operatorname{sh}\phi + (2 - \sigma)(1 - \operatorname{ch}\phi)(R - S)]}{S(R^2 - S^2)},$$

$$C_n = \frac{2\sigma(i + j)\operatorname{sh}\phi}{R - S} + \frac{2m[\sigma\phi S\operatorname{sh}\phi + (2 - \sigma)(1 - \operatorname{ch}\phi)(R - S)]}{\phi(R^2 - S^2)},$$

$$D_n = \frac{2\sigma(i + j)(1 - \operatorname{ch}\phi)}{R - S} + \frac{2m[(2 - \sigma)(R - S)\operatorname{sh}\phi + \sigma\phi(R - S\operatorname{ch}\phi)]}{\phi(R^2 - S^2)},$$

where $R = (3 + \sigma)\operatorname{sh}\phi$ and $S = (1 - \sigma)\phi$. These values substituted in (58) constitute the solution for W_{00} for $y = 0, b$ free. If $\lambda = 1$ and $\alpha = \beta = 0$, this solution reduces to that published by D. L. Holl.[13]

CASE III. *The edges $y = 0, b$ are clamped.* The boundary conditions are $W_{00} = \partial W_{00}/\partial y = 0$. Equation (58) and necessary Fourier expansions combined with these edge conditions yield

$$B_n = -2(i + j),$$

$$A_n\operatorname{sh}\phi + B_n\operatorname{ch}\phi + C_n\phi\operatorname{sh}\phi + D_n\phi\operatorname{ch}\phi = -2(i + j + m),$$

$$A_n\phi + D_n\phi = -2m,$$

$$A_n\phi\operatorname{ch}\phi + B_n\phi\operatorname{sh}\phi + C_n\phi(\phi\operatorname{ch}\phi + \operatorname{sh}\phi) + D_n\phi(\phi\operatorname{sh}\phi + \operatorname{ch}\phi) = -2m.$$

[12] S. Iguchi, *Eine Lösung für die Berechnung der biegsamen rechteckigen Platten*, Julius Springer, Berlin, 1933.

[13] D. L. Holl, *The deflection of an isotropic rectangular plate under the action of continuous and concentrated loads when supported at two opposite edges*, Iowa State College Journal of Science, vol. 9 (1935), pp. 597-607.

The solution of this system is

$$A_n = -\frac{2(i+j)(1-\operatorname{ch}\phi)}{\phi+\operatorname{sh}\phi} - \frac{2m\phi(1-\operatorname{ch}\phi)}{\phi^2-\operatorname{sh}^2\phi},$$

$$B_n = -2(i+j),$$

$$C_n = \frac{2(i+j)\operatorname{sh}\phi}{\phi+\operatorname{sh}\phi} + \frac{2m[\phi^2\operatorname{sh}\phi+(1-\operatorname{ch}\phi)(\phi+\operatorname{sh}\phi)]}{\phi(\phi^2-\operatorname{sh}^2\phi)},$$

$$D_n = \frac{2(i+j)(1-\operatorname{ch}\phi)}{\phi+\operatorname{sh}\phi} + \frac{2m(\operatorname{sh}^2\phi-\phi^2\operatorname{ch}\phi)}{\phi(\phi^2-\operatorname{sh}^2\phi)}.$$

These values substituted in equation (58) constitute the solution for W_{00} for $y = 0, b$ clamped. If $\lambda = 1$ and $\alpha = \beta = 0$, this solution reduces to that given by Holl.[14]

c. **Solution for** W_{20}. With W_{00} known it is now possible to solve for W_{20}. Since W_{20} is a bi-harmonic function of the second order of magnitude as compared with W_{00}, a solution of the following form is assumed.

$$\begin{aligned} W_{20} \doteq {} & \frac{-PAh^2}{2D}\left[\left(\lambda+\frac{\beta y}{b}\right)(x^2-ax)+\frac{\alpha}{3a}(x^3-a^2x)\right] \\ & -\frac{2Ph^2}{aD}\sum_n \frac{\sin\theta x}{\theta^3}[F_n\operatorname{ch}\theta y+G_n\operatorname{sh}\theta y+H_n\theta y\operatorname{ch}\theta y+I_n\theta y\operatorname{sh}\theta y]. \end{aligned} \tag{59}$$

Pinned edge condition (50.1) at $x = 0, a$ requires that $A = \dfrac{-(8-3\sigma)}{10(1-\sigma)}$. F_n, G_n, H_n, and I_n will now be determined by the conditions at $y = 0, b$ for all three cases.

CASE I. *The edges $y = 0, b$ are pinned.* The edge conditions are $W_{20} = 0$ and (51.1). The results are

$$F_n = -\frac{(8-3\sigma)}{10(1-\sigma)}(i+j),$$

$$G_n = -\frac{(8-3\sigma)}{10(1-\sigma)}\cdot\frac{[(i+j)(1-\operatorname{ch}\phi)+m]}{\operatorname{sh}\phi},$$

$$H_n = I_n = 0.$$

When these values are substituted in equation (59), the solution for W_{20}, for $y = 0, b$ pinned, results. It is easily shown that W_{20} may be written in the form

$$W_{20} = -\frac{(8-3\sigma)h^2}{10(1-\sigma)}\nabla^2 W_{00}. \tag{60}$$

[14] Loc. cit.

CASE II. *The edges $y = 0, b$ are free.* The edge conditions are (51.1) and (52.1). The results are

$$F_n = \frac{-(8+\sigma)}{10(1-\sigma)} C_n + \frac{1}{5} B_n + \frac{8m(1-\operatorname{ch}\phi)}{5S(R+S)},$$

$$G_n = \frac{-(8+\sigma)}{10(1-\sigma)} D_n + \frac{1}{5} A_n + \frac{4m(2\operatorname{sh}\phi + S)}{5S(R+S)},$$

$$H_n = \frac{1}{5} D_n - \frac{4m \operatorname{sh}\phi}{5\phi(R+S)},$$

$$I_n = \frac{1}{5} C_n - \frac{4m(1-\operatorname{ch}\phi)}{5\phi(R+S)},$$

where A_n, B_n, C_n and D_n are the constants determined for W_{00} for CASE II. W_{20} may be written in the form

$$W_{20} = \frac{-(8+\sigma)h^2}{10(1-\sigma)} \nabla^2 W_{00} - \frac{2h^2}{5} \frac{\partial^2 W_{00}}{\partial x^2} + \frac{4Ph^2}{5(1-\sigma)aD} \sum_n \frac{\sin \theta x}{\theta^3}$$

$$\Big[i + j + \frac{my}{b} + \frac{2m}{\phi(R+S)} \{ -2(1-\operatorname{ch}\phi)\operatorname{ch}\theta y - (2\operatorname{sh}\phi + S)\operatorname{sh}\theta y + (1-\sigma)\theta y \operatorname{sh}\phi \operatorname{ch}\theta y + (1-\sigma)\theta y(1-\operatorname{ch}\phi)\operatorname{sh}\theta y \} \Big]. \tag{61}$$

CASE III. *The edges $y = 0, b$ are clamped.* The edge conditions are $W_{20} = \frac{\partial W_{20}}{\partial y} = 0$. The results are

$$F_n = \frac{8-3\sigma}{20(1-\sigma)} B_n, \qquad G_n = \frac{8-3\sigma}{20(1-\sigma)} A_n,$$

$$H_n = \frac{8-3\sigma}{20(1-\sigma)} D_n, \qquad I_n = \frac{8-3\sigma}{20(1-\sigma)} C_n.$$

By use of these values solution (59) may be written as

$$W_{20} = \frac{-(8-3\sigma)h^2}{10(1-\sigma)} \frac{\partial^2 W_{00}}{\partial x^2}. \tag{62}$$

When $\lambda = 1$ and $\alpha = \beta = 0$, all three of these solutions for W_{20} reduce to those given by Garabedian.[15]

d. **Solution for** $W_{2n,0}$ ($n \geqq 2$). Now it is possible to solve for W_{40}. Since W_{40} is a harmonic function of the fourth order of magnitude as compared with W_{00}, a solution of the following form is assumed.

$$W_{40} = \frac{AP}{D}\left(\lambda + \frac{\alpha x}{a} + \frac{\beta y}{b}\right) + \frac{Ph^4}{aD} \sum_n \frac{\sin \theta x}{\theta} [J_n \operatorname{ch}\theta y + K_n \operatorname{sh}\theta y]. \tag{63}$$

[15] Footnote 4, vol. 195.

The boundary conditions at $x = 0, a$ are $W_{40} = 0$ and (56.2) which reduces to $\frac{\partial^2 W_{40}}{\partial x^2} = 0$. These conditions require that $A = 0$. J_n and K_n will now be determined by the conditions at $y = 0, b$ for each of the three cases.

CASE I. The boundary conditions at $y = 0, b$ are $W_{40} = 0$ and (57.2). These conditions require that $J_n = K_n = 0$. Therefore $W_{40} = 0$. By inspection of equations (16) it readily follows that $W_{2n,0} = 0$ $(n \geqq 2)$.

CASE II. The free edge conditions are (57.2) and (58.2). They require that

$$J_n = \frac{2(8+\sigma)}{5(1-\sigma)} I_n, \qquad K_n = \frac{2(8+\sigma)}{5(1-\sigma)} H_n.$$

With these values solution (63) may be written

$$W_{40} = \frac{-(8+\sigma)h^2}{10(1-\sigma)}\left[\nabla^2 W_{20} - \frac{8-3\sigma}{10(1-\sigma)} Ph^2\left(\lambda + \frac{\alpha x}{a} + \frac{\beta y}{b}\right)\right]. \tag{64}$$

By a consideration of orders of magnitude or by (16) it is evident that $W_{2n,0} = 0$ $(n > 2)$.

CASE III. The clamped edge conditions are $W_{40} = \frac{\partial W_{40}}{\partial y} = 0$. These conditions are satisfied when $J_n = K_n = 0$. Hence $W_{40} = 0$. As before, it follows that $W_{2n,0} = 0$ $(n \geqq 2)$.

All of these solutions for W_{40} reduce to those given by Garabedian[16] for the case $\lambda = 1$ and $\alpha = \beta = 0$.

e. **Solution for U_0 and V_0.** It remains to determine U_0 and V_0 to complete the problem. It has been pointed out that W_{01} and $\frac{\partial U_{00}}{\partial x} + \frac{\partial V_{00}}{\partial y}$ are of the same order of magnitude as the load which is of the fourth order of magnitude as compared with W_{00}. Therefore

$$\frac{\partial U_{00}}{\partial x} + \frac{\partial V_{00}}{\partial y} = a + bx + cy$$
$$+ \sum_n \frac{\sin\theta x}{a\theta}[F_n \operatorname{ch}\theta y + G_n \operatorname{sh}\theta y + H_n\theta y \operatorname{ch}\theta y + I_n \theta y \operatorname{sh}\theta y].$$

However equation (41) requires that this expression be a harmonic function. Therefore $H_n = I_n = 0$. With this background one may consider the following forms of solutions.

$$U_{00} = c_0 + c_1 x + c_2 y + c_3 x^2 + c_4 xy + c_5 y^2 + \sum_n \frac{\cos\theta x}{a\theta^2}[A_n \operatorname{ch}\theta y + B_n \operatorname{sh}\theta y],$$

$$V_{00} = k_0 + k_1 x + k_2 y + k_3 x^2 + k_4 xy + k_5 y^2 + \sum_n \frac{\sin\theta x}{a\theta^2}[C_n \operatorname{ch}\theta y + D_n \operatorname{sh}\theta y].$$

[16] See footnote 15.

It is obvious from the definitions of a pinned edge and a free edge as given in this paper that $T = 0$ is the only restriction on U_0 and V_0 for Cases I and II. Since this condition is not sufficient to determine all the constants in the above general forms, the solutions obtained here must be recognized as satisfactory solutions and not necessarily the unique solutions. Garabedian[17] avoided this difficulty in his solution for a uniform load by assuming that U_{00} and V_{00} were closed linear functions of x, y. However, this assumption does not permit a satisfactory solution in the pinned-clamped case as it requires that T_1 be different from zero at the pinned edges $x = 0, a$. Love (p. 462) adds the condition $S = 0$, but this condition does not yield a reasonable solution. Furthermore S. Woinowsky-Krieger[18] has shown that S is different from zero in general.

Equation (48.0) is the only condition on U_{00} and V_{00} at $x = 0, a$. It gives

$$c_1 + \sigma k_2 = \frac{\sigma\lambda P}{4\mu}, \qquad c_4 + 2\sigma k_5 = \frac{\sigma\beta P}{4\mu b}, \qquad 2c_3 + \sigma k_4 = \frac{\sigma\alpha P}{4\mu a}.$$

Therefore choose

$$c_0 = c_2 = c_5 = k_0 = k_1 = k_3 = 0;$$

$$c_1 = k_2 = \frac{\sigma\lambda P}{2E}; \qquad c_4 = 2k_5 = \frac{\sigma\beta P}{2Eb}; \qquad 2c_3 = k_4 = \frac{\sigma\alpha P}{2Ea},$$

where the quantity E is Young's modulus. Equation (49.0) at $y = 0, b$ yields $D_n = \sigma A_n$ and $C_n = \sigma B_n$. If the plate be fixed in space by making $V_{00} = 0$ at $y = 0$ and $U_{00} = 0$ at $x = 0$, then $B_n = A_n = 0$. Consequently $D_n = C_n = 0$ and the final solution may be written

$$U_{00} = \frac{\sigma x P}{2E}\left(\lambda + \frac{\alpha x}{2a} + \frac{\beta y}{b}\right), \tag{65}$$

$$V_{00} = \frac{\sigma y P}{2E}\left(\lambda + \frac{\alpha x}{a} + \frac{\beta y}{2b}\right). \tag{66}$$

Since U_{20} and V_{20} are of the second order of magnitude as compared to U_{00} and V_{00}, they must either be zero or constants. But $V_{20} = 0$ at $y = 0$ and $U_{20} = 0$ at $x = 0$ to fix the plate in space. Therefore they are zero everywhere. Consequently $U_{2n,0} = V_{2n,0} = 0$ $(n \geqq 1)$. This means that (65) and (66) give the complete horizontal displacement of the middle surface. Although not a unique solution it is a rational one in the sense that it yields a displacement at the middle surface which is one-half of what the displacement would be if the plate were resting on a complete foundation. If $\lambda = 1$ and $\alpha = \beta = 0$, these solutions reduce to those published by Garabedian[19] for a uniform load.

[17] See footnote 15.

[18] S. Woinowsky-Krieger, *Der Spannungszustand in dicken elastischen Platten*, Ingenieur-Archiv, vol. 4 (1933), pp. 203-226, 305-331.

[19] See footnote 15.

The values of U_{00} and V_{00} for the pinned-clamped case are most easily obtained by imposing the conditions $U_{00} = V_{00} = 0$ at $y = 0, b$ first and then requiring that $T_1 = 0$ at $x = 0, a$. The final results are

$$U_{00} = \frac{\sigma P}{2E}\Bigg[\lambda x - \frac{a}{6}(3\lambda + \alpha) - \frac{a\beta y}{2b} + \frac{\alpha x^2}{2a} + \frac{xy\beta}{b} + \sum_n \frac{2\cos\theta x}{a\theta^2}\left\{(i+j)\operatorname{ch}\theta y + \frac{(i+j)(1-\operatorname{ch}\phi) + m}{\operatorname{sh}\phi}\operatorname{sh}\theta y\right\}\Bigg], \tag{67}$$

$$V_{00} = \frac{\sigma P}{2E}\left[\lambda y + \frac{\alpha xy}{a} + \frac{\beta y^2}{2b} - \sum_n \frac{\sin\theta x}{a\theta^2}\left\{\frac{2(i+j)+m}{\operatorname{sh}\phi}\phi\operatorname{sh}\theta y\right\}\right]. \tag{68}$$

It readily follows that $U_{2n,0} = V_{2n,0} = 0$ $(n \geqq 1)$.

Woinowsky-Krieger's[20] solution for an infinite plate strip pinned at the edges $x = 0, a$ is readily obtained from any one of the three cases solved here. It is only necessary to set $V_0 = 0$ and redetermine U_0 to satisfy the conditions at the edges $x = 0, a$. Then in the limit as $b \to \infty$ the solutions given here reduce to those Woinowsky-Krieger gives for the infinite plate strip.

IOWA STATE COLLEGE.

[20] Loc. cit.

ON MATRICES OF INTEGERS AND COMBINATORIAL TOPOLOGY

By Hassler Whitney

1. **Introduction.** Our object is to give an elementary account of some algebraic theorems, with some immediate applications in combinatorial topology, in particular, in the theory of homology and cohomology[1] groups. The theorems are to a certain extent known, if in somewhat different forms.

The main tool in the algebraic part is the theory of group pairs, and in particular the question of when one group "resolves" or "completely resolves" another. The main theorems are on the existence of extensions of a homomorphism, and the existence of solutions of linear equations, with a matrix of integers and elements of an abelian group as unknowns. In each theorem, two types of conditions are employed, one using mod m properties, the other using group pairs. We shall use only discrete groups (except in Theorem 1).

The applications to topology are concerned with the relations between the ordinary homology theory, and the newer "dual" theory. An illustration of the convenience of the newer theory is given in Appendix I. For other illustrations, we mention the duality theorems (Kolmogoroff, Alexander, Čech, etc.), and properties of maps (see a following paper).

I. Group pairs and homomorphisms

2. **Group-pairs.** All groups will be abelian. 0 will denote the identity in any group. Let

$$mG = \text{all elements } mg,\ g \text{ in } G\ (m \text{ an integer}),$$

$$_mG = \text{all } g \text{ in } G \text{ such that } mg = 0.$$

Then $0G$ contains 0 alone. $G - G'$ is the difference (factor) group of G over the subgroup G'.

Let G, H, Z be three groups. If to each g in G and h in H there corresponds a $z = gh$ in Z, and both distributive laws are satisfied, we say G and H form a *group*

Received December 1, 1936; presented to the American Mathematical Society December, 1936.

[1] The relation of "coboundary" is becoming of increasing importance. Other terms have been used: dual boundary, upper boundary, inverse boundary, boundary in the dual subdivision, derived (of a function). The present term (accepted by E. Čech) offers definite advantages. In differential geometry, the (exact alternating) covariant tensors and the contravariant differentials play the same rôle as cocycles and (contra)cycles in the combinatorial theory; in fact, they may be obtained directly by a passage to the limit. Moreover, the prefix co is very convenient to handle.

pair with respect to Z. If $gh = 0$, we say g and h are *orthogonal*. For any subgroup H' of H, set

$$(G, H') = \textit{nullifier}^2 \text{ of } H' \text{ in } G = \text{all elements } g \text{ in } G \text{ orthogonal to every } h \text{ in } H'.$$

Define (H, G') similarly.

If $(G, H') = 0$, we say H' *resolves* G.[3] If $(G, {}_mH') = mG$, we say H' *m-resolves* G.[4] Clearly "H' 0-resolves G" and "H' resolves G" are equivalent statements. If H' m-resolves G for all integers $m \geqq 0$, we say H' *resolves G completely*. Note that

$$\text{(2.1)} \qquad \text{if } H' \text{ resolves } G, \text{ then } (G, mH') = {}_mG \qquad \text{(all } m).$$

(For $m = 1$, this is the definition.) For if $mg \neq 0$, choose h in H' so that $(mg)h \neq 0$; then $g(mh) \neq 0$.

Let G and Z be groups. A *Z-character* of G is a homomorphism of G into (part of) Z.[5] Let $Ch_Z(G)$ be the group of Z-characters of G. Given g in G and ϕ in $Ch_Z(G)$, set $g \cdot \phi = \phi(g)$; then these groups form a pair. Clearly *any group resolves any of its character groups*. The converse is false.

G is *infinitely divisible* if $mG = G$ for each m.

3. **Examples.** Let I_0 be the group of integers, and $I_\mu = I_0 - \mu I_0$ the group of integers mod μ. Using ordinary multiplication into $I_{(\mu,\nu)}$, it is seen that I_μ resolves I_ν if and only if μ is a multiple of ν. I_0 m-resolves no group G unless $m = 0$ or 1 or $mG = G$. The general question of when I_μ m-resolves I_ν is rather complicated. However, *I_μ completely resolves I_μ* if $\mu > 0$ and ordinary multiplication mod μ is used. This is a consequence of Theorem 1 below.[6] If R_1 is the group of the real (or the rational) numbers mod 1, then I_0 and R_1 completely resolve each other, as is easily seen; this also follows from Theorem 1.

R_1 is infinitely divisible; no group with a finite number of generators is.

THEOREM 1. *A discrete or compact (or locally bicompact[7]) group completely resolves and is completely resolved by its R_1-character group.*

[2] The term annihilator has been used, but this word seems unnecessarily barbarous.

[3] If H' resolves G, and $g_1 \neq g_2$, then $(g_1 - g_2)h \neq 0$ for some h in H', and $g_1h \neq g_2h$. Thus distinct elements of G may be shown to be distinct by multiplying by elements of H'.

[4] Thus, by multiplying by elements of H', we can tell whether a given element of G is divisible by m or not. Any H' 1-resolves G.

[5] It is often important to introduce a topology into groups and character groups; we shall not do this here.

[6] Direct proof: Note that

$$mI_\mu = \text{all numbers in } I_\mu \equiv 0 \bmod (m, \mu),$$
$${}_mI_\mu = \text{all numbers in } I_\mu \equiv 0 \bmod \mu/(m, \mu).$$

Suppose $ab \equiv 0 \bmod \mu$ for all b in ${}_mI_\mu$. Taking $b = \mu/(m, \mu)$ shows that $a \equiv 0 \bmod (m, \mu)$, and a is in mI_μ.

[7] See L. Pontrjagin, Annals of Math., vol. 35 (1934), pp. 361–388; also van Kampen, Annals of Math., vol. 36 (1935), pp. 448–463.

That G and $H = Ch_{R_1}(G)$ resolve each other is shown by Pontrjagin, loc. cit., Theorem 5 and Remark 4. Now set $H' = mH$. As H resolves G, $(G, mH) = {}_mG$, by (2.1). By Pontrjagin, Theorem 4, $(H, (G, mH)) = mH$; hence $(H, {}_mG) = mH$, and G m-resolves H. As m is arbitrary, G completely resolves H. Similarly H completely resolves G.

Note that a group may resolve and be resolved by its character group, and not resolve it completely. For an example, take $G = H = Z = I_0$.

4. Extensions of homomorphisms. Suppose part of a group is mapped into another group. Can the map be extended through the first group so that it is a homomorphism? We shall give two answers to the question.

THEOREM 2.[8] *Let A be a free group with a finite number of generators, let A' be a subset of A, and let f map the elements of A' into the group G. The definition of f may be extended over A so that it is a homomorphism if and only if* (a) *holds:*

(a) *For any elements $a_1, \cdots, a_s$ of A' and integers $m \geqq 0, \alpha_1, \cdots, \alpha_s$,*

$$\sum \alpha_i a_i \text{ in } mA \text{ implies } \sum \alpha_i f(a_i) \text{ in } mG. \tag{4.1}$$

Suppose that A and H, and G and H, form group pairs, and let H resolve G completely. If the following condition holds, the extension of f over A is possible:

(b) *For any elements $a_1, \cdots, a_s$ of A', integers $\alpha_1, \cdots, \alpha_s$, and element h of H,*

$$(\sum \alpha_i a_i)h = 0 \text{ implies } (\sum \alpha_i f(a_i))h = 0. \tag{4.2}$$

The necessity of (a) is clear:

$$\sum \alpha_i f(a_i) = f(\sum \alpha_i a_i) = f(ma) = mf(a).$$

We turn to the sufficiency. Let A^* be the subgroup of A generated by the elements of A'. For each $a = \sum \alpha_i a_i$ in A^*, set

$$f(a) = \sum \alpha_i f(a_i).$$

To prove that f is uniquely determined in A^*, suppose

$$\sum \alpha_i a_i = \sum \beta_i a_i, \qquad \sum (\beta_i - \alpha_i)a_i = 0.$$

[8] The theorem will be strengthened in Appendix II. It is easily seen that the conditions may be weakened as follows: (a) and (b) may be replaced by the corresponding pairs of hypotheses

(a_1) $\Sigma\alpha_i a_i = 0$ implies $\Sigma\alpha_i f(a_i) = 0$,

(a_2) a in mA implies $f(a)$ in mG;

(b_1) $\Sigma\alpha_i a_i = 0$ implies $(\Sigma\alpha_i f(a_i))h = 0$ (all h),

(b_2) $ah = 0$ (any fixed h, a in A') implies $f(a)h = 0$.

Further, the values $m = 0, 1$ may be omitted in (a). We note from the proof that (b) *implies* (a).

The theorems in Alexandroff-Hopf, *Topologie*, I, pp. 591-3, are consequences of the strengthened theorem (using footnote 6). Our proof has naturally much in common with theirs.

(We may carry out both sums over the same terms.) Suppose (a) holds. Then as 0 is in $0A$,

$$\sum (\beta_i - \alpha_i)f(a_i) = \sum \beta_i f(a_i) - \sum \alpha_i f(a_i)$$

is in $0B$, hence it is 0, and the last two terms are equal. Now suppose (b) holds. Then for any h in H,

$$[\sum (\beta_i - \alpha_i)a_i]h = 0;$$

hence

$$[\sum \beta_i f(a_i) - \sum \alpha_i f(a_i)]h = 0,$$

and as H resolves G, the term in brackets equals 0.

Obviously f is a homomorphism in A^*. Moreover, as each linear combination of elements of A^* is a linear combination of elements of A', (a) or (b) is easily seen to hold in A^* if it holds in A'.

Let $b_1, \cdots, b_r, b_{r+1}, \cdots, b_s$ form a base for A, and let $m_1, \cdots, m_r$ be positive integers such that $a_1 = m_1b_1, \cdots, a_r = m_rb_r$ form a base for A^*.[9] Suppose (a) holds. As a_i is in m_iA, $f(a_i)$ is in m_iG, and we may choose g_i so that $m_ig_i = f(a_i)$ $(i = 1, \cdots, r)$. If (b) holds, $a_ih = b_i(m_ih) = 0$ for all h in ${}_{m_i}H$; hence $f(a_i)h = 0$ for all h in ${}_{m_i}H$, and as H resolves G completely, $f(a_i)$ is in m_iG. Again g_i exists. Set

$$f(b_i) = g_i \ (i = 1, \cdots, r), \quad \text{and} \quad f(b_i) = 0 \ (i = r + 1, \cdots, s).$$

The resulting homomorphism of A into G obviously agrees with the one already found in A^*.

Remarks. If A' is a subgroup of A with division (i.e., ma in A', $m \neq 0$, implies a in A'), and f is a homomorphism of A' into G, we may always extend f over A. For we may choose a base $b_1, \cdots, b_r, b_{r+1}, \cdots, b_s$ for A such that $b_1, \cdots, b_r$ form a base for A', and set $f(b_i) = 0$ $(i > r)$. Further, any homomorphism of any subgroup A' of any group A into G may be extended over A, if G is infinitely divisible; see Pontrjagin, loc. cit., proof of Lemma 1.

II. Linear equations with integer coefficients

5. Cycles, etc., of a matrix. Let η be a matrix of integers:

$$\eta = \begin{Vmatrix} \eta_1^1 & \cdots & \eta_n^1 \\ \cdots & \cdots & \cdots \\ \eta_1^p & \cdots & \eta_n^p \end{Vmatrix}. \tag{5.1}$$

Given a group X, let $X^s = X + \cdots + X$, a direct sum with s terms. The elements of X^s may be written $(x_1, \cdots, x_s)$ or $(x^1, \cdots, x^s)$. We define the *boundary* and *coboundary* (with respect to η) of elements of X^n and X^p by

$$\partial(x^1, \cdots, x^n) = (\sum \eta_j^1 x^j, \cdots, \sum \eta_j^p x^j) \text{ in } X^p, \tag{5.2}$$

$$\delta(x_1, \cdots, x_p) = (\sum \eta_1^j x_j, \cdots, \sum \eta_n^j x_j) \text{ in } X^n. \tag{5.3}$$

[9] See Alexandroff-Hopf, loc. cit., p. 568, no. 24.

Any element (5.2) we call an *X-boundary*; these form a group $\mathbf{B}^X(\eta)$. Define similarly *X-coboundaries* and $\mathbf{B}_X(\eta)$.

Any element of $X^n[X^p]$ whose boundary [coboundary] vanishes we shall call an *X-cycle* [*X-cocycle*]. We form with these the groups $\mathbf{C}^X(\eta)$ and $\mathbf{C}_X(\eta)$.

Given a group pair G, H, we form a group pair G^s, H^s by

$$(5.4) \qquad (g^1, \cdots, g^s)\cdot(h_1, \cdots, h_s) = \sum g^i h_i .$$

THEOREM 3. *If H resolves G, or G resolves H, then*

$$(5.5) \qquad \mathbf{C}^G(\eta) = (G^n, \mathbf{B}_H(\eta)), \textit{ or } \mathbf{C}_H(\eta) = (H^p, \mathbf{B}^G(\eta)).$$

We shall prove the first; the other is obtained by transposing η. For any G-cycle $(g^1, \cdots, g^n)$ and H-coboundary $\delta(h_1, \cdots, h_p)$,

$$(5.6) \quad (g^1, \cdots, g^n)\cdot\delta(h_1, \cdots, h_p) = \sum_i g^i \sum_j \eta^j_i h_j = \partial(g^1, \cdots, g^n)\cdot(h_1, \cdots, h_p) = 0.$$

Now suppose $(g^1, \cdots, g^n)$ satisfies this relation for all $(h_1, \cdots, h_p)$. Taking all the $h_k = 0$ but the j-th gives $(\sum \eta^j_i g^i)h = 0$ (all h); as H resolves G, $\sum \eta^j_i g^i = 0$ (all j), and $(g^1, \cdots, g^n)$ is a G-cycle.

6. Linear equations in integers.

We shall now prove

THEOREM 4.[10] *Let η and G be given; if we use (β), assume H resolves G completely. $(g^1, \cdots, g^p)$ is a G-boundary, i.e., the equations*

$$(6.1) \qquad \sum_{j=1}^{n} \eta^i_j \bar{g}^j = g^i \qquad (i = 1, \cdots, p)$$

are solvable for $\bar{g}^1, \cdots, \bar{g}^n$, if and only if one of the following is true:

(α) *For any integers $m \geqq 0, \alpha_1, \cdots, \alpha_p$,*

$$(6.2) \qquad \sum_{i=1}^{p} \alpha_i \eta^i_j \equiv 0 \text{ mod } m \textit{ (all } j\textit{) implies } \sum_{i=1}^{p} \alpha_i g^i \textit{ is in } mG;$$

in other words, every I_0-cocycle mod m[11] *(or, every I_m-cocycle) is orthogonal* mod m *to $(g^1, \cdots, g^p)$.*

(β) *For any $(h_1, \cdots, h_p)$ in H^p,*

$$(6.3) \qquad \sum_{i=1}^{p} \eta^i_j h_i = 0 \textit{ (all } j\textit{) implies } \sum_{i=1}^{p} g^i h_i = 0;$$

in other words, every H-cocycle is orthogonal to $(g^1, \cdots, g^p)$.

Example. Let $G = I_\mu$, μ a prime; then η may be considered as a matrix of integers mod μ. Using footnote 7, both (α) and (β) (with $H = Z = I_\mu$) reduce to testing (α) for the single integer $m = \mu$.

[10] A proof of the first half, (α), of this theorem, using a standard theorem on linear equations (see Veblen, *Analysis Situs*, Second edition, Appendix 2) has been furnished to me by H. T. Engstrom.

[11] x is a cocycle mod m if its coboundary is divisible by m; it is orthogonal mod m to y if xy is divisible by m. We might state the condition as follows: $(g^1, \cdots, g^m)$ is in the "complete nullifier" of $\mathbf{C}_{I_0}(\eta)$.

The necessity of either condition is trivial; we shall prove the sufficiency. Set $A = I_0^n =$ all $(a_1, \cdots, a_n)$ (a's integers). Let A' be the rows $\eta^1, \cdots, \eta^p$ of η, and set

$$f(\eta^i) = f(\eta_1^i, \cdots, \eta_n^i) = g^i \qquad (i = 1, \cdots, p). \tag{6.4}$$

Suppose f has been extended over A so that it is a homomorphism of A into G. Then set

$$\bar{g}^1 = f(1, 0, \cdots), \qquad \bar{g}^2 = f(0, 1, \cdots), \qquad \cdots. \tag{6.5}$$

Then

$$\sum_{j=1}^{n} \eta_j^i \bar{g}^j = \sum_{j=1}^{n} \eta_j^i f(\overset{1}{0}, \cdots, \overset{j}{1}, \cdots, \overset{n}{0}) = f(\eta_1^i, \cdots, \eta_n^i) = g^i.$$

Thus we need merely show that (a) or (b) of Theorem 2 is satisfied.

Suppose first (α) holds, and $\sum \alpha_i \eta^i$ is in mA. Then $\sum \alpha_i \eta_j^i \equiv 0 \bmod m$ (all j), hence $\sum \alpha_i g^i = \sum \alpha_i f(\eta^i)$ is in mG, and (a) holds.

Suppose next (β) holds. Let A and H form a group pair by setting $(a_1, \cdots, a_n)h = (a_1 h, \cdots, a_n h)$. Suppose

$$(\sum \alpha_i \eta^i)h = \sum \alpha_i(\eta_1^i h, \cdots, \eta_n^i h) = 0.$$

Then

$$\sum \alpha_i \eta_j^i h = \sum \eta_j^i h_i = 0 \qquad \text{(all } j\text{)},$$

where $h_i = \alpha_i h$. Hence, by (β),

$$(\sum \alpha_i f(\eta^i))h = \sum f(\eta^i)h_i = \sum g^i h_i = 0,$$

and (b) holds. This completes the proof.

Remark. (β) of this theorem and a similar theorem with η transposed say that

$$\mathbf{B}^G(\eta) = (G^p, \mathbf{C}_H(\eta)) \text{ if } H \text{ resolves } G \text{ completely,} \tag{6.6}$$

$$\mathbf{B}_H(\eta) = (H^n, \mathbf{C}^G(\eta)) \text{ if } G \text{ resolves } H \text{ completely.} \tag{6.7}$$

III. Applications to topology

7. Cycles, boundaries, cocycles, and coboundaries. Let K be a (finite) complex, with cells σ_i^r $(i = 1, \cdots, \alpha^r)$ of dimension r, and incidence matrices ${}^r\partial = \| {}^r\partial_j^i \|$. The *boundary* and *coboundary* of σ_i^r are

$$\partial\sigma_i^r = \sum_j {}^r\partial_i^j \sigma_j^{r-1}, \qquad \delta\sigma_i^r = \sum_j {}^{r+1}\partial_j^i \sigma_j^{r+1}. \tag{7.1}$$

In terms of the coefficients of chains, boundaries and coboundaries are given by (5.2) and (5.3) with η_j^i replaced by ${}^r\partial_j^i$ and ${}^{r+1}\partial_j^i$.

An *r-X-chain* is a linear form $\sum x_i\sigma_i^r$ (x_i in X); we may consider it as an element of X^{α^r}. Define the groups of

$$\begin{array}{ll} r\text{-}X\text{-}cycles: & {}^r\mathbf{C}^X = \mathbf{C}^X({}^r\partial), \\ r\text{-}X\text{-}boundaries: & {}^r\mathbf{B}^X = \mathbf{B}^X({}^{r+1}\partial), \\ r\text{-}X\text{-}cocycles: & {}^r\mathbf{C}_X = \mathbf{C}_X({}^{r+1}\partial), \\ r\text{-}X\text{-}coboundaries: & {}^r\mathbf{B}_X = \mathbf{B}_X({}^r\partial). \end{array}$$

Then the *r-X-homology groups* and *r-X-cohomology groups* are the difference groups

$$(7.2)\qquad {}^r\mathbf{H}^X = {}^r\mathbf{C}^X - {}^r\mathbf{B}^X, \qquad {}^r\mathbf{H}_X = {}^r\mathbf{C}_X - {}^r\mathbf{B}_X.$$

If G and H form a group pair, then r-G-chains and r-H-chains may be multiplied, using (5.4). By (5.6),

$$(7.3)\qquad \partial A\cdot B = A\cdot\delta B$$

for any $(r+1)$-chain A and r-chain B.

Equations (5.5), (6.6) and (6.7) give

$$(7.4)\qquad {}^r\mathbf{C}^G = (G^{\alpha^r}, {}^r\mathbf{B}_H) \qquad \text{if } H \text{ resolves } G,$$

$$(7.5)\qquad {}^r\mathbf{C}_H = (H^{\alpha^r}, {}^r\mathbf{B}^G) \qquad \text{if } G \text{ resolves } H,$$

$$(7.6)\qquad {}^r\mathbf{B}^G = (G^{\alpha^r}, {}^r\mathbf{C}_H) \qquad \text{if } H \text{ resolves } G \text{ completely},$$

$$(7.7)\qquad {}^r\mathbf{B}_H = (H^{\alpha^r}, {}^r\mathbf{C}^G) \qquad \text{if } G \text{ resolves } H \text{ completely}.$$

Expressed in words, we have (using (7.3))

THEOREM 5. *With suitable G and H, an r-chain (with coefficients in G or H) is an r-G-cycle, or an r-G-boundary, or an r-H-cocycle, or an r-H-coboundary, if and only if it is orthogonal to every r-H-coboundary, or r-H-cocycle, or r-G-boundary, or r-G-cycle.*

THEOREM 6. *An r-G-chain is a cycle, or a cocycle, if and only if it is orthogonal to every I_0-coboundary, or I_0-boundary. It is a boundary, or a coboundary, if and only if it is orthogonal* mod m *to every I_m-cocycle, or I_m-cycle, for all m ($\neq 1$).*

The proof of the first half is like that of Theorem 3; the second half follows from Theorem 4, (α).

8. **Homology and cohomology groups.** We shall find a case in which the homology groups are determined by the cohomology groups or vice versa.

THEOREM 7. *For any G and H, the groups ${}^r\mathbf{H}^G$, ${}^r\mathbf{H}_H$ form a pair, if G and H form a pair, as follows. Given elements ξ of ${}^r\mathbf{H}^G$ and ζ of ${}^r\mathbf{H}_H$, choose a cycle C of ξ and a cocycle D of ζ, and set $\xi\zeta = CD$.*

Theorem 5 shows that the definition depends on ξ and ζ alone.

THEOREM 8. *Let Z be infinitely divisible. If $H = Ch_Z(G)$ and G resolves H completely, or if $G = Ch_Z(H)$ and H resolves G completely, then*

$$(8.1)\qquad {}^r\mathbf{H}_H = Ch_Z({}^r\mathbf{H}^G), \text{ or } {}^r\mathbf{H}^G = Ch_Z({}^r\mathbf{H}_H).$$

We shall prove the first equation; the proof of the second is the same. Each element ζ of ${}^r\mathbf{H}_H$ determines a Z-character ϕ_ζ of ${}^r\mathbf{H}^G$ by the last theorem. Suppose $\zeta \neq \zeta'$. Take cocycles D in ζ and D' in ζ'; then $D' - D$ is not a coboundary. By (7.7) there is a cycle C such that $C(D' - D) \neq 0$, $CD \neq CD'$. Let ξ be the homology class of C; then $\xi\zeta \neq \xi\zeta'$, and $\phi_\zeta \neq \phi_{\zeta'}$. Therefore ${}^r\mathbf{H}_H$ determines in a (1-1) way a subgroup of the Z-characters of ${}^r\mathbf{H}^G$.

To show that this is the whole group, take any Z-character ϕ of ${}^r\mathbf{H}^G$. For any cycle C with homology class ξ, set $\psi(C) = \phi(\xi)$; this is a Z-character of ${}^r\mathbf{C}^G$. By the remark following Theorem 2, we may extend it to a Z-character ψ of G^{α^r}, the group of all r-G-chains. Set

$$\psi_i(g) = \psi(0, \cdots, g, \cdots, 0) \qquad (g \text{ in the } i\text{-th place}).$$

As $H = Ch_Z(G)$, there is an h_i such that $gh_i = \psi_i(g)$ (all g). Set

$$D = (h_1, \cdots, h_{\alpha^r}) = \sum h_i \sigma_i^r;$$

then for any r-G-chain $C = (g^1, \cdots, g^{\alpha^r})$,

$$CD = \sum g^i h_i = \sum \psi_i(g^i) = \psi(g^1, \cdots, g^{\alpha^r}) = \psi(C).$$

As ψ maps all boundaries into 0, D is a cocycle, by (7.5); let ζ be its homology class. Then for any ξ in ${}^r\mathbf{H}^G$, $\xi\zeta = \phi(\xi)$. This completes the proof.

Corollary. ${}^r\mathbf{H}_{I_0}$ *is isomorphic to the direct sum of the reduced r-th homology group and the* $(r-1)$*-th torsion group, all with integer coefficients.*

This follows from the last theorem, Theorem 1, and Alexandroff-Hopf, *Topologie* I, p. 234, equation (17′). A direct proof is not hard to give, using the ordinary elementary divisor theory.

Remarks. The theorem does not hold for arbitrary Z. For, consider the projective plane, using $G = R_1$, $Z = I_2$; then $H = 0$, and hence ${}^2\mathbf{H}_H = 0$. But ${}^2\mathbf{H}^G$ = torsion group of dimension 1 $= I_2$ (see Alexandroff-Hopf, loc. cit., p. 234), and hence $Ch_Z({}^2\mathbf{H}^G) = CH_{I_2}I_2 = I_2$.

Given H, set $G = Ch_{R_1}(H)$; then $H = Ch_{R_1}(G)$, and G completely resolves H, by Theorem 1. Hence ${}^r\mathbf{H}_H = Ch_{R_1}({}^r\mathbf{H}^G)$. As ${}^r\mathbf{H}^G$ is a topological invariant, ${}^r\mathbf{H}_H$ is a topological invariant of K.

Appendix I

On closed complexes and pseudomanifolds

For the general theory, see Alexandroff-Hopf, *Topologie* I, Ch. VII, §1. We refer to this work as AH. K will always denote a homogeneous[12] n-complex, and σ_1, σ_2, $\cdots$, its n-cells. The chain σ_i means the n-chain $\sum \delta_{ij}\sigma_j$ with coefficient 1 for σ_i and 0 for all other σ_j. Write $C \backsim D$ if C is cohomologous to D.

[12] That is, each k-cell ($k < n$) is on some n-cell.

9. **Closed complexes.** We say K is *closed* if no σ_i is $\backsim 0$ in K. That this definition agrees with the ordinary one is shown by Theorem 9 below.

LEMMA. *σ_i is not $\backsim 0$ if and only if it is contained in some cycle C* mod m *for some $m \neq 1$.*[13]

By Theorem 6, σ_i is not $\backsim 0$ if and only if for some $m \neq 1$ there is a cycle C mod m such that $C \cdot \sigma_i \not\equiv 0$ mod m. Write $C = a\sigma_i + \cdots$; then $C \cdot \sigma_i = a$, and the above holds if and only if $a \not\equiv 0$ (mod m), i.e., if and only if C contains σ_i.

THEOREM 9. *K is closed if and only if each σ_i is contained in a cycle* mod m *for an $m \neq 1$.*

This is a consequence of the lemma. Alexandroff-Hopf, p. 275, Satz Ia, shows that our definition is equivalent to the ordinary one.

10. **Irreducibly closed complexes.** We say K is *irreducibly closed* if it is closed but no proper subcomplex is.

THEOREM 10.[14] *If K is irreducibly closed, then*

(a) *the cohomology group ${}^n\mathbf{H}_{I_0}$ is cyclic* (*of order $\neq 1$*), *and*

(b) *no σ_i is $\backsim 0$, but each σ_i is $\backsim$ some multiple of each σ_j.*

To prove (b) we shall show first that $\sigma_i \backsim 0$ in $K - \sigma_j$. If not, then by the lemma σ_i is contained in a cycle C mod m, $m \neq 1$, C in $K - \sigma_j$. Let K' be the complex containing the cells of C; then each of its cells is contained in C, and hence K' is closed, by Theorem 9; but K' is in $K - \sigma_j$, a contradiction. Say $\sigma_i = \delta A$ in $K - \sigma_j$; then

$$\delta A = \sigma_i - p_{ij}\sigma_j \text{ in } K \qquad \text{(some } p_{ij}\text{)},$$

and $\sigma_i \backsim p_{ij}\sigma_j$, as required.

To prove (a), take a σ_i, and let m be the smallest positive integer such that $m\sigma_i \backsim 0$; if there is none, set $m = 0$. Each $\mu\sigma_i$ is a cocycle (as $\dim(K) = n$), and determines an element of ${}^n\mathbf{H}_{I_0}$. Further, given any element ξ of ${}^n\mathbf{H}_{I_0}$, determined by the cocycle C,

$$C = \sum_j \alpha_j \sigma_j \backsim \sum_j \alpha_j(p_{ji}\sigma_i) = \mu\sigma_i,$$

and ξ is determined by $\mu\sigma_i$. Hence ${}^n\mathbf{H}_{I_0}$ is cyclic, of order m.

Remark. There exist complexes in which some p_{ij} is $\neq \pm 1$; see AH, p. 280, Bemerkung.

THEOREM 11.[15] *For the homogeneous n-complex K to be irreducibly closed, either of the two following conditions is* (*necessary and*) *sufficient.*

(α) *This is* (b) *of Theorem* 10.

(β) ${}^n\mathbf{H}_{I_0}(K) \neq 0$, *but* ${}^n\mathbf{H}_{I_0}(K') = 0$ *for any proper subcomplex K' of K.*

[13] The condition $m \neq 1$ could be omitted; for a cycle mod 1 contains no cells, i.e., is $\equiv 0$ mod 1.

[14] Compare AH, p. 277, Satz IV and Satz V.

[15] See AH, Theorems on p. 280.

Suppose (α) holds; then K is closed. If K is not irreducibly closed, then there is a proper (homogeneous) subcomplex K' which is closed. Take σ_i in K', σ_j in $K - K'$, and say

$$\sigma_i - a\sigma_j = \delta A \text{ in } K.$$

Write

$$A = A_1 + A_2, \qquad A_1 \text{ in } K', \qquad A_2 \text{ in } K - K';$$

we shall show that $\sigma_i = \delta A_1$ in K'. As no σ_k^{n-1} of A_2 is on a cell of K', δA_2 has no cells in K'; hence the part of δA in K' is the part of δA_1 in K', and this is σ_i. But then $\sigma_i \backsim 0$ in K', contradicting the assumption that K' is closed.

Suppose (β) holds. Let D be a cocycle not $\backsim 0$. By Theorem 6, there is a cycle C mod m for some $m \neq 1$ with $C \cdot D \neq 0$; hence $C \neq 0$. The cells contained in C form a closed complex K', by Theorem 9. Now let K^* be any proper subcomplex of K. Then ${}^n\mathbf{H}_{I_0}(K^*) = 0$; consequently each σ_i of K^* is $\backsim 0$ in K^*, and K^* is not closed. Therefore $K' = K$, and K is irreducibly closed.

11. **Pseudomanifolds.** A *pseudomanifold* is a strongly connected[16] homogeneous complex in which each $(n - 1)$-cell is on either one or two n-cells; it is *closed* if each $(n - 1)$-cell is on two n-cells.

THEOREM 12.[17] *If the pseudomanifold K is not closed, then ${}^n\mathbf{H}_{I_0}$ vanishes; if it is closed, then it is an irreducibly closed complex, and ${}^n\mathbf{H}_{I_0}$ is cyclic of order* 0 *or* 2 *according as K is orientable or not.*

Take any succession of distinct n-cells $\sigma_0^n, \sigma_1^n, \cdots, \sigma_s^n$, σ_i^n and σ_{i+1}^n having the common $(n - 1)$-face σ_i^{n-1}. We may orient the cells so that σ_i^{n-1} is on σ_i^n negatively and on σ_{i+1}^n positively. Let A be the sum of these σ_i^{n-1}; then

$$\delta A = \sigma_s^n - \sigma_0^n.$$

It follows that any two n-cells of K, with the proper orientations, are $\backsim$.

If K is not closed (as a pseudomanifold), there is an $(n - 1)$-cell σ^{n-1} on just one n-cell σ^n; then $\delta\sigma^{n-1} = \pm\sigma^n$, and $\sigma^n \backsim 0$; hence each $\sigma_i^n \backsim 0$, and ${}^n\mathbf{H}_{I_0}$ vanishes. Suppose K is closed. Then each $\delta\sigma^{n-1}$ and hence each δA^{n-1} is a chain the sum of whose coefficients is even, and hence cannot be any σ_i^n; as each $\sigma_i \backsim \pm$ each σ_j, K is an irreducibly closed complex, by Theorem 11, (α).

If K is not orientable, a cell σ_0^n may be joined to itself by a succession of n- and $(n - 1)$-cells so as to reverse the orientation. Let A be the sum of the properly oriented $(n - 1)$-cells of the chain; then $\delta A = 2\sigma_0^n$, so that $2\sigma_0^n$ and hence any $2\sigma_i^n$ is $\backsim 0$; hence ${}^n\mathbf{H}_{I_0}$ is of order 2. Otherwise, we may orient the cells of K concurrently; then each $\delta\sigma_i^{n-1}$ and hence each δA^{n-1} is a chain, the sum of whose coefficients vanishes, and no $m\sigma_i^n$ is $\backsim 0$ for $m \neq 0$. It follows that ${}^n\mathbf{H}_{I_0}$ is the infinite cyclic group.

[16] That is, each two n-cells are joined by a succession of n- and $(n - 1)$-cells.

[17] Compare AH, p. 281, Satz VIII.

APPENDIX II

We shall show that Theorem 2 holds if A is any group with a finite number of generators. Extend f over A^* as before. Let $b^1, \cdots, b^r$ form a base for A, and $a^1, \cdots, a^s$, a base for A^*. Say

$$a^i = \sum_{j=1}^{r} \eta_j^i b^j \qquad (i = 1, \cdots, s).$$

Let μ_i be the order of b^i ($\mu_i = 0$ if b^i is of infinite order). Set

$$\eta_j^{s+i} = \mu_i \text{ if } j = i, \; = 0 \text{ if } j \neq i \qquad (i, j = 1, \cdots, r);$$

then

$$\sum_{j=1}^{r} \eta_j^{s+i} b^j = \mu_i b^i = 0 \qquad (i = 1, \cdots, r).$$

Set

$$g^i = f(a^i) \qquad (i = 1, \cdots, s), \qquad g^{s+i} = 0 \qquad (i = 1 \cdots, r).$$

Suppose $\sum_{i=1}^{s+r} \alpha_i \eta_j^i = mk_j$ (all j). Then

$$\sum_{i=1}^{s} \alpha_i a^i = \sum_{i=1}^{s} \sum_{j=1}^{r} \alpha_i \eta_j^i b^j = \sum_{i=1}^{s+r} \sum_{j=1}^{r} \alpha_i \eta_j^i b^j = m \sum_{j=1}^{r} k_j b^j$$

is in mA, and hence, using (a) or (b) (see footnote 8),

$$f\left(\sum_{i=1}^{s} \alpha_i a^i\right) = \sum_{i=1}^{s} \alpha_i f(a^i) = \sum_{i=1}^{s+r} \alpha_i g^i$$

is in mG. Therefore, by Theorem 4, there are elements $\bar{g}^1, \cdots, \bar{g}^r$ of G such that

$$\sum_{j=1}^{r} \eta_j^i \bar{g}^j = g^i \qquad (i = 1, \cdots, s + r).$$

Setting $f'(\sum \alpha_i b^i) = \sum \alpha_i \bar{g}^i$ defines uniquely a homomorphism of A into G; for

$$f'(\mu_i b^i) = \sum_{j=1}^{r} \eta_j^{s+i} \bar{g}^j = g^{s+i} = 0 \qquad (i = 1, \cdots, r).$$

Further,

$$f'(a^i) = f'(\sum \eta_j^i b^j) = \sum \eta_j^i \bar{g}^j = g^i = f(a^i) \qquad (i = 1, \cdots, s),$$

so that f' is an extension of f.

HARVARD UNIVERSITY.

ON THE MAPS OF AN n-SPHERE INTO ANOTHER n-SPHERE

By Hassler Whitney

1. **Introduction.** It is well known that to each map[1] f of an n-sphere S^n into another one S_0^n ($n \geqq 1$ always) there corresponds a number d_f, the *degree* of f, and $d_f = d_g$ if f and g are homotopic (see §2). H. Hopf[2] has proved the converse theorem, that if $d_f = d_g$, then f and g are homotopic. The object of this note is to give an elementary proof of the latter theorem. The methods will be used and extended in later papers.

In an appendix we give somewhat briefly a proof of the theorem for the case that $d_f = 0$. This is the only case needed in the following paper; the general theorem then follows from that paper. The second proof is more intuitive geometrically than the first, but complete details would make it perhaps more lengthy.

2. **On deformations.** A *deformation* of one space S in another S_0 is a family $\phi_t(p)$ ($0 \leqq t \leqq 1$, p in S) of maps of S into S_0, continuous in both variables together. Given maps f and g of S into S_0, if there exists a deformation ϕ_t such that $\phi_0 \equiv f$ and $\phi_1 \equiv g$, we say f and g are *homotopic*. If f is homotopic to g, where $g(p) \equiv P_0$ (all p in S), we say f is *homotopic to zero*, and f may be *shrunk to the point* P_0.

Suppose S and S_0 are complexes, K_0 is a simplicial subdivision of S_0, and f maps S into S_0. Then, for a sufficiently fine simplicial subdivision K of S, the following is true. To each vertex V of K we may choose a vertex $g(V)$ of a cell of K_0 which contains $f(V)$, so that the vertices of any cell of K go into the vertices of a cell of K_0. This determines uniquely a "simplicial map" g of K into K_0, affine in each cell (see §5); moreover, f is homotopic to g.

3. **The degree of a map.** Let S_0^n be the unit n-sphere in $(n + 1)$-space, let K_0^n be a simplicial triangulation of S_0^n, and let σ_0^n be an n-cell of K_0^n. We choose K_0^n so that if P_1 is a point of σ_0^n and P_0 is the antipodal point of S_0^n, each great semicircle from P_1 to P_0 intersects the boundary $\partial\sigma_0^n$ of σ_0^n in exactly one point. By pushing along these semicircles, we define a deformation Ω_t of the identity $\Omega_0(p) \equiv p$ into a map Ω_1, where $\Omega_1(p) \equiv P_0$ for p in $S_0^n - \sigma_0^n$.

Let σ^k be a k-cell ($k \leqq n$), in fixed correspondence with a k-simplex, and let

Received December 1, 1936; presented to the American Mathematical Society, October, 1936.

[1] All maps will be assumed continuous.

[2] See Alexandroff-Hopf, *Topologie*, I, Berlin, 1935, pp. 501–505. See also the reference to Lefschetz in the following paper.

f map σ^k into S_0^n. We say f is *standard* if $f(p) \equiv P_0$, or, $k = n$ and for some affine map ϕ of $\sigma^k = \sigma^n$ into σ_0^n, $f(p) \equiv \Omega_1(\phi(p))$. In any case, $f(p) \equiv P_0$ in $\partial\sigma^k$.[3] The map f of an n-complex K^n into S_0^n is *standard* if it is standard over each k-cell ($k \leqq n$).

We may orient S_0^n by orienting σ_0^n. Let K^n be a simplicial triangulation of the oriented n-sphere S^n, and let f be a standard map of K^n into S_0^n. Let σ^n be an (oriented) n-cell of K^n. If $f(p) \equiv P_0$ in σ^n, we set $d_f(\sigma^n) = 0$. Otherwise, there is a simplicial map ϕ of σ^n into (the whole of) σ_0^n such that $f(p) \equiv \Omega_1(\phi(p))$ in σ^n; we set $d_f(\sigma^n) = 1$ or -1 according as ϕ is positive or negative. We define the *degree* of f by

$$d_f = \sum_{\sigma^n} d_f(\sigma^n). \tag{3.1}$$

4. **The theorem.** In homology theory it is shown how to attach to each map f of S^n into S_0^n (both spheres oriented) an integer d_f, the degree of the map. Moreover, if f is homotopic to g, then $d_f = d_g$, and if S^n and S_0^n are triangulated and f is standard, then d_f is given by (3.1).

Suppose f and g map S^n into S_0^n, and $d_f = d_g$. Then for a sufficiently fine subdivision K^n of S^n, both f and g can be deformed into simplicial maps and hence into standard maps ϕ and ψ. As f and ϕ, also g and ψ, are homotopic, $d_\phi = d_\psi$. By Theorem 1 below, ϕ is homotopic to ψ; hence f is homotopic to g. Therefore this theorem furnishes the converse of the statements above.

THEOREM 1. *If ϕ and ψ are standard maps of S^n into S_0^n, using the same subdivision K^n of S^n, and $d_\phi = d_\psi$, then ϕ is homotopic to ψ.*

From the proof below, the following corollary is apparent.

COROLLARY. *If $\phi(V) = \psi(V) = P_0$ for a fixed vertex V of K^n, we can make V remain at P_0 throughout the deformation.*

In fact, if $n \geqq 2$, all vertices of K^n remain at P_0. If $n = 1$, we may choose the chains of cells in §8 so that in no chain do we pass over V; then V is never moved.

THEOREM 2. *For any integer γ there is a map f of S^n into S_0^n with $d_f = \gamma$.*

To prove this, subdivide S^n into $\alpha \geqq |\gamma|$ n-cells. Let ϕ map $|\gamma|$ of these cells simplicially into σ_0^n, positively or negatively according as $\gamma > 0$ or $\gamma < 0$ (if $\gamma \neq 0$), and set $f(p) = \Omega_1(\phi(p))$ in these cells and $f(p) = P_0$ elsewhere. Clearly $d_f = \gamma$. Note that the degree of the identity map of S_0^n into itself is 1.

The remainder of the paper is devoted to the proof of Theorem 1.

5. **Coördinates p_t in a cell.** Any simplicial complex K^n is homeomorphic to a complex $\bar{K}^n$ in euclidean space whose cells are straight. Using $\bar{K}^n$, we define straightness in K^n, the center of a cell (i.e., center of mass of its vertices), etc. Hence an "affine map" of one cell into another has meaning. Let σ be a cell of K^n, and a, the center of σ. For each point p of the boundary $\partial\sigma$ of σ let p_t be the point of the segment ap such that $ap_t/ap = t$.

[3] There are $(n+1)!$ standard maps ϕ of σ^n into S_0^n with $\phi(p) \not\equiv P_0$.

6. **Certain deformations of simplexes.** We prove first a combinatorial lemma, needed in Lemma 2.

LEMMA 1. *Any even permutation of the letters $a_0a_1 \cdots a_n$ $(n \geqq 2)$ may be made by means of a succession of cyclic permutations, each on three of the letters.*

This is clear if $n = 2$; then any even permutation is cyclic. Suppose $n > 2$, and let $B = a_{\alpha_0} \cdots a_{\alpha_n}$ be any even permutation. If $\alpha_n \neq n$, bring a_{α_n} to the right end by a cyclic permutation; bring $a_{\alpha_{n-1}}$ next to a_{α_n}. Suppose $\alpha_0 \neq 0$. We then perform the two cyclic permutations

$$a_0 \cdots a_{\alpha_0} \cdots a_{\alpha_{n-1}} a_{\alpha_n} \to a_{\alpha_0} \cdots a_{\alpha_{n-1}} \cdots a_0 a_{\alpha_n} \to a_{\alpha_0} \cdots a_0 \cdots a_{\alpha_n} a_{\alpha_{n-1}} .$$

If $n \geqq 4$ and a_{α_1} is not now in the second place, we perform two cyclic permutations to bring it there, again interchanging $a_{\alpha_{n-1}}$ and a_{α_n}, etc. When $a_{\alpha_0}, \cdots, a_{\alpha_{n-3}}$ are in their correct places, $a_{\alpha_{n-2}}$ is also; as B is even and the above permutations are even, $a_{\alpha_{n-1}}$ and a_{α_n} are also in their correct positions.

LEMMA 2. *Let $\sigma^n = a_0 \cdots a_n$ be a simplex, and let $a_{\alpha_0} \cdots a_{\alpha_n}$ be an even permutation of its vertices. Then there is a deformation ϕ_t of σ^n in itself, such that $\phi_0(p) \equiv p$, $\phi_1(a_s) = a_{\alpha_s}$, ϕ_1 is affine, and ϕ_t for each t is a homeomorphism both in σ^n and in its boundary.*

If $n = 0$ or 1, the lemma is trivial. Suppose that $n = 2$; say $a_{\alpha_0}a_{\alpha_1}a_{\alpha_2} = a_1a_2a_0$. Let $\phi_t(a_i)$ be the point p of a_ia_{i+1} (setting $2 + 1 = 0$) for which $a_ip/a_ia_{i+1} = t$. Let ϕ_t map the segment a_ia_{i+1} into the broken line $\phi_t(a_i)a_{i+1}\phi_t(a_{i+1})$ so that, if the line were straightened, the map would be linear. For any point p_u (see §5) interior to σ^2, set $\phi_t(p_u) = (\phi_t(p))_u$. As $\phi_t(a) = a =$ center of mass of σ^2, ϕ_1 is easily seen to be affine.

Now suppose $n > 2$; consider first a cyclic permutation, changing say $a_0a_1a_2$ into $a_1a_2a_0$. Set $\sigma = a_0a_1a_2$, $\sigma' = a_3 \cdots a_n$, and let $[p, q, u]$ for p in σ, q in σ', $0 \leqq u \leqq 1$, be the point r of the segment pq for which $pr/pq = u$. Define ϕ_t in σ as above. For any point $[p, q, u]$ not in σ, set $\phi_t[p, q, u] = [\phi_t(p), q, u]$. We show that ϕ_t is a homeomorphism. Suppose $\phi_t[p, q, u] = \phi_t[p', q', u']$; then $[\phi_t(p), q, u] = [\phi_t(p'), q', u']$, which implies $\phi_t(p) = \phi_t(p')$, $q = q'$, $u = u'$;[4] as ϕ_t is a homeomorphism in σ, $p = p'$ also. Further, given $[p, q, u]$ and t, we may find a p^* for which $\phi_t(p^*) = p$; then $\phi_t[p^*, q, u] = [p, q, u]$. The other properties of ϕ_t are clear, and the lemma for this case is proved. Now take any permutation. We may obtain it by cyclic permutations as in Lemma 2; the corresponding deformations together give the required deformation.

7. **Two types of deformations of S^n in S_0^n.** Let ϕ' be a standard map of S^n into S_0^n, and let σ and σ' be oriented n-cells of K^n with the common $(n - 1)$-face τ:

$$\sigma = a_0a_1 \cdots a_n , \qquad \sigma' = -a_0'a_1 \cdots a_n , \qquad \tau = a_1 \cdots a_n .$$

(a) Suppose $d_{\phi'}(\sigma) = 1$, $d_{\phi'}(\sigma') = 0$; we shall deform ϕ' into ϕ'' so that $d_{\phi''}(\sigma) = 0$, $d_{\phi''}(\sigma') = 1$, leaving $K^n - (\sigma + \sigma')$ fixed.

(b) Suppose $d_{\phi'}(\sigma) = 1$, $d_{\phi'}(\sigma') = -1$; we shall obtain $d_{\phi''}(\sigma) = d_{\phi''}(\sigma') = 0$.

[4] This is so if $0 < u < 1$, as we may assume.

In each case ϕ'' will be a standard map.

(a) Set $\sigma_1 = a_0 a_2 \cdots a_n$, $\sigma_1' = a_0' a_2 \cdots a_n$, or if $n = 1$, then $\sigma_1 = a_0$, $\sigma_1' = a_0'$. Let θ_1 and θ_2 be the affine maps of σ_1 into τ and σ_1' determined by sending a_0 into a_1 and a_0' respectively. For each p in σ_1, let $\alpha(p, u)$ run linearly along the segments $p\theta_1(p)$ and $\theta_1(p)\theta_2(p)$ as u runs from 0 to 1 and from 1 to 2. Set

$$\phi_t'[\alpha(p, u)] = \begin{cases} \phi'[\alpha(p, u - t)] & (t \leqq u), \\ \phi'[\alpha(p, 0)] & (t > u), \end{cases} \tag{7.1}$$

and $\phi_t'(p) = \phi'(p)$ in $K^n - (\sigma + \sigma')$. As $\phi'(p) \equiv P_0$ in $\partial\sigma_1 + \partial\sigma_2$, this is clearly a deformation of $\phi' = \phi_0$ into a map $\phi'' = \phi_1$. The map ϕ'' in σ' is obtained from the map ϕ' in σ by replacing $a_0, a_1, \cdots, a_n$ (which form $+\sigma$) by $a_1, a_0', \cdots, a_n$ (which form $+\sigma'$); hence $d_{\phi''}(\sigma') = d_{\phi'}(\sigma)$. Also $d_{\phi''}(\sigma) = 0$ as $\phi''(p) \equiv P_0$ in σ, and (a) is proved.

(b) Let λ and λ' be the affine maps of σ and σ' into σ_0^n such that $\phi'(p) = \Omega_1(\lambda(p))$ in σ and $= \Omega_1(\lambda'(p))$ in σ'. Say $\sigma_0^n = b_0 \cdots b_n$,

$$\lambda(a_i) = b_{k_i}, \qquad \text{and } \lambda'(a_0') = b_{l_0}, \qquad \lambda'(a_i) = b_{l_i} \qquad (i > 0).$$

As $d_{\phi'}(\sigma') = -d_{\phi'}(-\sigma')$, and hence

$$d_{\phi'}(\sigma) = d_{\phi'}(a_0 a_1 \cdots a_n) = -d_{\phi'}(\sigma') = d_{\phi'}(a_0' a_1 \cdots a_n),$$

$b_{l_0} \cdots b_{l_n}$ is an even permutation of $b_{k_0} \cdots b_{k_n}$. Applying Lemma 2, we find a deformation λ_t' of σ' in σ_0^n such that $\lambda_0' \equiv \lambda'$, λ_1' is affine, and

$$\lambda_1'(a_0') = \lambda(a_0), \qquad \lambda_1'(a_i) = \lambda(a_i) \qquad (i > 0). \tag{7.2}$$

Set

$$\phi_t'(p) = \begin{cases} \Omega_1(\lambda_t'(p)) & p \text{ in } \sigma', \\ \phi'(p), & p \text{ in } K^n - \sigma'. \end{cases} \tag{7.3}$$

Then as $\Omega_1(\lambda_t'(p)) \equiv P_0$ in $\partial\sigma'$, ϕ_t' is a deformation of ϕ' into a map $\phi^* \equiv \phi_1'$.

For each p in τ, let $\beta(p, u)$ be the point q of the segment $a_0 p$ of σ such that $a_0 q / a_0 p = u$, and let $\beta'(p, u)$ be the corresponding point of the segment $a_0' p$ in σ'. As λ and λ_1' are affine, (7.2) and (7.3) give

$$\phi^*[\beta(p, u)] = \phi^*[\beta'(p, u)] \qquad (p \text{ in } \tau, 0 \leqq u \leqq 1). \tag{7.4}$$

We deform ϕ^* into ϕ'' by setting

$$\phi_t^*[\beta(p, u)] = \phi_t^*[\beta'(p, u)] = \phi^*[\beta(p, (1 - t)u)], \tag{7.5}$$

and $\phi_t^*(p) = \phi^*(p)$ in $K^n - (\sigma + \sigma')$. This is clearly a deformation; (7.4) shows that $\phi_0^* \equiv \phi^*$. As $\phi''(p) \equiv \phi_1^*(p) \equiv P_0$ in $\sigma + \sigma'$, $d_{\phi''}(\sigma) = d_{\phi''}(\sigma') = 0$.

8. **Proof of Theorem** 1. Suppose there are cells of K^n mapped positively over S_0^n by ϕ, and also cells mapped negatively. Then we can find a chain $\sigma_0, \cdots, \sigma_\nu$ of adjacent n-cells of K^n such that

$$d_\phi(\sigma_0) = 1, \quad d_\phi(\sigma_1) = 0, \quad \cdots, \quad d_\phi(\sigma_{\nu-1}) = 0, \quad d_\phi(\sigma_\nu) = -1.$$

Using (a), §7, we deform ϕ in $\sigma_0 + \sigma_1$, then in $\sigma_1 + \sigma_2$, etc.; then, using (b), §7, we deform the map in $\sigma_{\nu-1} + \sigma_\nu$. The new map ϕ' has $d_{\phi'}(\sigma_i) = 0$ ($i = 0, \cdots, \nu$). Continue in this manner till no cells are mapped positively or none are mapped negatively over S_0^n; for definiteness, say the latter holds. Do the same for ψ. The new maps ϕ^* and ψ^* each have exactly $d_\phi = d_\psi$ cells mapped positively over S_0^n.

Suppose $d_{\phi^*}(\sigma) \neq d_{\psi^*}(\sigma)$ for some σ. Then let $\sigma_0, \sigma_1, \cdots, \sigma_\nu$ be a chain of adjacent n-cells such that

$$d_{\phi^*}(\sigma_0) = d_{\psi^*}(\sigma_\nu) = 1, \qquad d_{\phi^*}(\sigma_\nu) = d_{\psi^*}(\sigma_0) = 0,$$
$$d_{\phi^*}(\sigma_i) = d_{\psi^*}(\sigma_i) \qquad (0 < i < \nu).$$

Let $\sigma_0, \sigma_{k_1}, \cdots, \sigma_{k_s}$ be the cells of the chain for which $d_{\phi^*} = 1$. Using (a), §7, we deform ϕ^* over $\sigma_{k_s} + \sigma_{k_s+1}$ etc. until we have $d_{\phi_1^*}(\sigma_{k_s}) = 0$, $d_{\phi_1^*}(\sigma_\nu) = 1$; another succession of deformations makes $d_{\phi_2^*}(\sigma_{k_{s-1}}) = 0$, $d_{\phi_2^*}(\sigma_{k_s}) = 1$, etc. Finally $d_{\phi^{**}}(\sigma_0) = 0$, $d_{\phi^{**}}(\sigma_{k_i}) = 1$ (all i), and $d_{\phi^{**}}(\sigma_\nu) = 1$. $d_{\phi^{**}}(\sigma)$ differs from $d_{\psi^*}(\sigma)$ over fewer cells than $d_{\phi^*}(\sigma)$. Continuing in this manner, we deform ϕ^* into a map ϕ' with $d_{\phi'}(\sigma) = d_{\psi^*}(\sigma)$, all σ. ϕ' and ψ^* are standard. Applying Lemma 2, we deform ϕ' over each n-cell where necessary, to obtain ψ^*. (Compare the first half of the proof of (b), §7.) This completes the proof.

Appendix[5]

Let f be a map of S^n into S_0^n with the degree 0. We first deform it into a simplicial map and then into a standard map ϕ (see §§ 2, 3). To shrink ϕ to a point is equivalent to extending ϕ through the interior R of S^n (see the following paper, § 4). Let $\sigma_1, \cdots, \sigma_s$ and $\sigma_1', \cdots, \sigma_s'$ be the simplexes of S^n mapped positively and negatively over S_0^n respectively. Let T_i be a tube joining σ_i to σ_i' inside R. We may choose these so no two intersect, and also (to prove the corollary) so no one cuts the radius of R to the vertex V. Let $a_0 \cdots a_n$ and $a_0' \cdots a_n'$ be positive and negative orientations of σ_i and σ_i' respectively, such that $\lambda(a_i) = b_i$ and $\lambda'(a_i') = b_i$ determine simplicial maps of σ_i and σ_i' into σ_0^n, which in turn determine ϕ in σ_i and σ_i'. Now carry σ_i through T_i to σ_i', turning it so that a_i goes into a_i'; let $g_t(\sigma_i)$ be the position of σ_i after the time t. We do this so that $g_t(\sigma_i)$ does not intersect $g_{t'}(\sigma_i)$ if $t \neq t'$. (We are using a deformation theorem on simplexes in euclidean space, similar to but simpler than Lemma 2.) The definition of ϕ in R is as follows. For p not in any $g_t(\sigma_i)$, set $\phi(p) = P_0$. For p in $g_t(\sigma_i)$, choose q in σ_i so that $p = g_t(q)$, and set $\phi(p) = \phi(q)$.

Harvard University.

[5] Added in proof.

THE MAPS OF AN n-COMPLEX INTO AN n-SPHERE

By Hassler Whitney

1. **Introduction.** The classes of maps of an n-complex into an n-sphere were classified by H. Hopf[1] in 1932. Recently, W. Hurewicz[2] has extended the theorem by replacing the n-sphere by much more general spaces. Freudenthal[3] and Steenrod[4] have noted that the theorem and proof are simplified by using real numbers reduced mod 1 in place of integers as coefficients in the chains considered. We shall give here a statement of the theorem which seems the most natural; the proof is quite simple. As in the original proof by Hopf, we shall base it on a more general extension theorem.

The fundamental tool of the paper is the relation of "coboundary";[5] it has come into prominence in the last few years.

In later papers we shall classify the maps of a 3-complex into a 2-sphere and of an n-complex into projective n-space.

I. Elementary facts

2. **Boundaries and coboundaries.** Let K be a complex, with oriented cells σ_i^r (not necessarily simplicial) of dimension r, $r = 0, \cdots, n$. Let $\partial_{ij}^r = 1, -1$, or 0 according as σ_i^{r-1} is positively, negatively, or not at all, on the boundary of σ_j^r. An r-chain C^r is a linear form $\Sigma \alpha_i \sigma_i^r$, the α_i being integers (or elements of an abelian group). The *boundary* (or *contraboundary*) and *coboundary* of C^r are defined by

$$(2.1)\qquad \partial\Big(\sum_i \alpha_i \sigma_i^r\Big) = \sum_{i,j} \alpha_i \partial_{ji}^r \sigma_j^{r-1}, \qquad \delta\Big(\sum_i \alpha_i \sigma_i^r\Big) = \sum_{i,j} \alpha_i \partial_{ij}^{r+1} \sigma_j^{r+1}.$$

As in the ordinary theory, we say C^r is a *cocycle* if its coboundary vanishes, and C^r is *cohomologous* to D^r, $C^r \backsim D^r$, if $C^r - D^r$ is a coboundary. The relation $\delta\delta C^r = 0$ (easily proved; equivalent to $\partial\partial C^r = 0$) says that every coboundary

Received December 1, 1936; presented to the International Topological Conference in Moscow, September 1935, and to the American Mathematical Society, October 1935.

[1] H. Hopf, Commentarii Mathematici Helvetici, vol. 5 (1932), pp. 39–54. See also Alexandroff-Hopf, *Topologie* I, Ch. XIII. A recent proof has been given by S. Lefschetz, Fund. Math., vol. 27 (1936), pp. 94–115. In Lemma 3 he gives a new proof of the theorem of the preceding paper; the author does not understand how the final map is made simplicial.

[2] W. Hurewicz, Proc. Kön. Akad. Wet. Amsterdam, vols. 38–39 (1935–36); in particular, vol. 39, pp. 117–126. The full paper will appear in the Annals of Math.

[3] H. Freudenthal, Compositio Math., vol. 2 (1935), footnote 8.

[4] Unpublished.

[5] This is discussed briefly in §2. For further details, see our paper *On matrices of integers*, pp. 35–45 of this volume of this Journal. We refer to this paper as I. The relation of Theorems 2, 3 and 4 to the theorems as stated by Hopf are made apparent by the theorems in I. *The present paper is independent of* I.

is a cocycle. Hence we may define the difference group of the group of r-cocycles over the group of r-boundaries, forming the r-th *cohomology group*.[6]

3. **Normal maps of cells into S_0^n.** Let S_0^n be the (oriented) unit n-sphere in $(n+1)$-space. Let f map the (oriented) n-cell σ^n into S_0^n. We say f is *normal* if $f(p) \equiv P_0$, a fixed point of S_0^n, for p in the boundary $\partial\sigma^n$ of σ^n. This is equivalent to identifying the points of $\partial\sigma^n$ in σ^n, forming an n-sphere S^n, and mapping this sphere into S_0^n. Hence we may define the degree[7] $d_f(\sigma^n)$. If f and g are normal in σ^n and $d_f(\sigma^n) = d_g(\sigma^n)$, then we may deform f into g, keeping $\partial\sigma^n$ at P_0, by II, corollary.

Any map f of σ^r into S_0^n, $r < n$, may be shrunk to P_0: we deform f into a simplicial map, and apply Ω_t (see II, §3). P_0 being assumed a vertex of K_0^n, if $\partial\sigma^r$ is at P_0 it remains there during the deformation.

If K is any complex, let K^r be the subcomplex of K containing all its cells of dimension $\leqq r$. The map f of K into S_0^n is *normal* if $f(p) \equiv P_0$ for p in K^{n-1}. Suppose σ^n or S^n is subdivided into cells σ_i^n, and f is a normal map of it into S_0^n. Then the $d_f(\sigma_i^n)$ are defined, and

$$d_f(\sigma^n) \text{ or } d_f(S^n) = \sum_i d_f(\sigma_i^n). \tag{3.1}$$

To show this, subdivide σ^n or S^n further, so that we can deform f into a simplicial map, and apply Ω_1 (see II, §3). The above quantities are unchanged, and (3.1) is now a consequence of II, (3.1).

4. **On deformations.** We shall need the following elementary results. Let $K \times I$ be the product of K and the unit interval I, consisting of all pairs (p, t), p in K, $0 \leqq t \leqq 1$. The deformation $\phi_t(p)$ of K in S_0^n is equivalent to the map $\Phi(p, t) = \phi_t(p)$ of $K \times I$ into S_0^n. Hence ϕ_0 is homotopic to ϕ_1 if and only if Φ, defined over $K \times 0 + K \times 1$, may be extended over $K \times I$.

Let f map the boundary $\partial\sigma^r$ of σ^r into S_0^n. Then f is homotopic to zero (in $\partial\sigma^r$) if and only if it may be extended through σ^r. For the deformation $f_t(p)$ (p in $\partial\sigma^r$) into $f_1(p) \equiv P$ is equivalent to the map $f(p_{1-t})$ (see II, §5) $= f_t(p)$ of σ^r into S_0^n.

LEMMA 1. *If $\phi \equiv \phi_0$ maps σ^n into S, and the deformation ϕ_t of ϕ is defined over $\partial\sigma^n$, then its definition may be extended over σ^n.*

We define ϕ_t in σ^n by

$$\phi_t(p_u) = \begin{cases} \phi(p_{(1+t)u}) & \left(0 \leqq t \leqq \dfrac{1}{u} - 1\right), \\ \phi_{t+1-\frac{1}{u}}(p) & \left(\dfrac{1}{u} - 1 \leqq t \leqq 1\right). \end{cases} \tag{4.1}$$

[6] This is the character group of the homology group with numbers mod 1 as coefficient group.

[7] See pp. 46–50 of this volume of this Journal; we refer to this paper as II.

LEMMA 2. *Any map ϕ of K into S_0^n may be deformed into a normal one; all cells already at P_0 we may keep fixed.*

We deform the map successively so that $K^0, K^1, \cdots, K^{n-1}$ are at P_0. Suppose K^{r-1} is at P_0 (if $0 < r < n$). As each $\partial\sigma^r$ is at P_0, we may deform each σ^r into P_0, keeping $\partial\sigma^r$ at P_0 (see §3). This deformation, defined over K^r, is extended over all $(r + 1)$-cells, $(r + 2)$-cells, etc., by Lemma 1. It is now defined over K, and K^r is at P_0.

5. **Parts of cocycles.** Let K' be a subcomplex of K. Any r-chain C of K may be written $C' + C''$, the coefficients of cells of $K - K'$[8] [of K'] being zero in C' [in C'']. We say C' is *part* of C. Clearly the chain C' in K' is part of a cocycle if and only if $\delta C'$ cobounds in $K - K'$, i.e., if and only if for some chain C'' in $K - K'$, $\delta C' = \delta C''$. The $(r + 1)$-chains are chains of K.

6. **The product** $K \times I$. We subdivide $K \times I$ (see §4) by means of all cells $\sigma_i^r \times I$ (σ_i^r in K). Orient the cells $\sigma_i^r \times 0$ and $\sigma_i^r \times 1$ like the σ_i^r, and orient each $(r + 1)$-cell $\sigma_i^r \times I$ so that $\sigma_i^r \times 1$ is on its boundary positively. Then

$$\delta(\sigma_i^r \times 0) = -\sigma_i^r \times I + \cdots, \qquad \delta(\sigma_i^r \times 1) = \sigma_i^r \times I + \cdots, \tag{6.1}$$

$$\delta(\sigma_i^r \times I) = -\sum_j \partial_{ij}^{r+1}(\sigma_j^{r+1} \times I). \tag{6.2}$$

To prove (6.2), say $\delta(\sigma_i^r \times I) = A_{ij}^{r+1}(\sigma_j^{r+1} \times I) + \cdots$. Then

$$\begin{aligned}\delta\delta(\sigma_i^r \times 1) &= \delta[(\sigma_i^r \times I) + \sum_j \partial_{ij}^{r+1}(\sigma_j^{r+1} \times 1)]\\ &= (A_{ij}^{r+1} + \partial_{ij}^{r+1})(\sigma_j^{r+1} \times I) + \cdots = 0,\end{aligned}$$

and $A_{ij}^{r+1} = -\partial_{ij}^{r+1}$. The first equation in (6.1) is clear for $r = 0$; it is proved in succession for $r = 1, 2, \cdots$ by considering the coefficient of $\sigma_j^r \times I$ in $\delta\delta(\sigma_i^{r-1} \times 0)$.

THEOREM 1. *Let C_0 and C_1 be n-chains in $K = K^n$, and let D_0 and D_1 be the corresponding chains in $K \times 0$ and $K \times 1$. Then $D_0 + D_1$ (as a chain in $K \times I$) is part of a cocycle if and only if $C_0 \backsim C_1$ in K.*

Say

$$C_0 = \sum a_i\sigma_i^n, \qquad C_1 = \sum b_i\sigma_i^n.$$

Consider any n-chain

$$D = D_0 + D_1 + \sum h_j(\sigma_j^{n-1} \times I); \tag{6.3}$$

then, by (6.1) and (6.2),

$$\begin{aligned}\delta D &= -\sum a_i(\sigma_i^n \times I) + \sum b_i(\sigma_i^n \times I) - \sum h_j\partial_{ji}^n(\sigma_i^n \times I)\\ &= \sum_i [b_i - a_i - \sum_j h_j\,\partial_{ji}^n](\sigma_i^n \times I).\end{aligned} \tag{6.4}$$

[8] $K - K'$ is in general not a subcomplex of K, i.e., is not closed in K.

Suppose $D_0 + D_1$ is part of a cocycle D; then (6.4) set $= 0$ gives

$$\delta\Big(\sum_j h_j\sigma_j^{n-1}\Big) = \sum_{i,j} h_j \partial_{ji}^n \sigma_i^n = \sum_i (b_i - a_i)\sigma_i^n = C_1 - C_0,$$

and $C_0 \backsim C_1$. Conversely, suppose $C_1 - C_0 = \delta(\sum h_j\sigma_j^{n-1})$; then the last set of equations shows that the bracket in (6.4) vanishes, and hence D, defined by (6.3), is a cocycle.

II. **The theorems**

7. The extension theorem. We shall prove

THEOREM 2. *Let f be a normal map of the subcomplex K' of $K = K^{n+1}$ into S_0^n. Then f can be extended over K if and only if the chain*

$$D' = \sum_{\sigma_i^n \text{ in } K'} d_f(\sigma_i^n)\sigma_i^n \tag{7.1}$$

in K' is part of a cocycle.

First suppose D' is part of a cocycle $D = \sum a_i\sigma_i^n$:

$$a_i = d_f(\sigma_i^n) \qquad (\sigma_i^n \text{ in } K'), \qquad \sum_i a_i\partial_{ij}^{n+1} = 0 \ \ (\text{all } j). \tag{7.2}$$

f maps $(K')^{n-1}$ into P_0; set $f(p) \equiv P_0$ in K^{n-1}. Let f map each σ_i^n not in K' into S_0^n with the degree a_i (see II, Theorem 2); then (7.2) holds for all σ_i^n. Consider any $(n+1)$-cell σ_j^{n+1} of $K - K'$. Using (3.1), we find

$$\begin{aligned} d_f(\partial\sigma_j^{n+1}) &= d_f\Big(\sum_i \partial_{ij}^{n+1}\sigma_i^n\Big) = \sum_i \partial_{ij}^{n+1} d_f(\sigma_i^n) \\ &= \sum_i \partial_{ij}^{n+1} a_i = 0. \end{aligned} \tag{7.3}$$

Hence f, considered only in $\partial\sigma_j^{n+1}$, is homotopic to zero (II, Theorem 1), and f may be extended over σ_j^{n+1} (see §4). Thus we extend f throughout K.

Now suppose f is extended throughout K. By Lemma 2, we deform f into a normal map, leaving $(K')^{n-1}$, and hence also K', fixed. Call the new map f again, and define the a_i and D by (7.2). Then D' is part of D. By §4, f, in each $\partial\sigma_j^{n+1}$, is homotopic to zero; hence (7.3) holds, and D is a cocycle.

Remark. If f is any map of K' into S_0^n, we may deform it into a normal map ϕ, by Lemma 2. From Lemma 1, it is apparent that f can be extended over K if and only if ϕ can be. Define D' by (7.1). By Theorem 2, $\delta D'$ has zero coefficients over cells of K', and is therefore a chain, which is clearly a cocycle, of $K'' = K - K'$. By Theorem 3, Remark, if f is also deformed into the normal map ψ, defining the chain C' of K', then $C' \backsim D'$ in K', and hence for some H in K',

$$C' - D' = (\delta H)' = \delta H - (\delta H)''.$$

Therefore $\delta C' - \delta D' = \delta[(\delta H)'']$, which lies in K''. Thus the cohomology class in K'' of $\delta D'$ is uniquely determined by f, and we have (using Theorem 2): *f may be extended over K if and only if its cohomology class thus defined in K'' is $\backsim 0$ in K''.*

8. **The classes of maps of K^n into S_0^n.** If we put two maps of K^n into S_0^n into the same class if they are homotopic, the maps fall into classes, the *homotopy classes.* To any normal map f of K^n into S_0^n we let correspond a chain C_f as in (7.1).

THEOREM 3. *The normal maps ϕ and ψ of $K = K^n$ into S_0^n are homotopic if and only if $C_\phi \backsim C_\psi$.*

Set $\Phi(p \times 0) = \phi(p), \Phi(p \times 1) = \psi(p)$; then ϕ is homotopic to ψ if and only if Φ may be extended through $K \times I$ (see §4). If D_0 and D_1 correspond to C_ϕ and C_ψ in $K \times 0$ and $K \times 1$, Theorem 2 shows that this is possible if and only if $D' = D_0 + D_1$ is part of a cocycle in $K \times I$. By Theorem 1, this is true if and only if $C_\phi \backsim C_\psi$.

Remark. If K is of any dimension and ϕ and ψ are homotopic, then C_ϕ and C_ψ are cocycles and $C_\phi \backsim C_\psi$. The first statement follows from Theorem 2; the second follows on considering ϕ and ψ in K^n alone.

THEOREM 4. *The classes of maps of K^n into S_0^n are in* $(1 - 1)$ *correspondence with the elements of the n-th cohomology group of K with integer coefficients. The correspondence is given by deforming the map f into a normal one and taking the cohomology class of the resulting cocycle. In particular, f is homotopic to zero if and only if the corresponding cohomology class is zero.*

The deformation is possible, by Lemma 2. The cohomology class is uniquely determined by f, and non-homotopic maps determine different classes, by Theorem 3. Finally, to each cohomology class corresponds a map; we take a cocycle C of the class, and let f map each σ^n normally into S_0^n with the degree equal to its coefficient in C (see II, Theorem 2).

9. **The Theorem of Hurewicz.** Let Q_0 be a fixed point of a space S. Then the classes of maps of S_0^r into S for which P_0 goes into Q_0 form an abelian group, the *r-th homotopy group* of S.[9] If f maps σ^n [or S_0^n] into S, and $f(p) \equiv Q_0$ in $\partial\sigma^n[f(P_0) = Q_0]$, we may call the corresponding homotopy element the *degree* $d_f(\sigma^n)[d_f(S_0^n)]$ of f. (If $S = S_0^n$, the n-th homotopy group is the group of integers, as was seen in II, so that this is a natural generalization of the term degree.) The fundamental formula (3.1) holds still. The theorems of the preceding paper become matters of definition. The proofs in the present paper hold without change, and we have a new version of the Theorem of Hurewicz:

THEOREM 5. *Theorems 2, 3 and 4 hold if we replace S_0^n by any locally contractible space S_0 whose r-th homotopy groups vanish for $r < n$, and replace the integers by the n-th homotopy group of S as coefficient group in the chains and cohomology classes.*

Hurewicz also shows that in the above space S_0 the n-th homotopy group is the same as the n-th homology group with integer coefficients.

HARVARD UNIVERSITY.

[9] See Hurewicz, loc. cit. We assume a knowledge of the fundamental properties of homotopy groups.

ON THE EXTREME POINTS OF CONVEX SETS

By G. Baley Price

Introduction. A convex set is a set such that if it contains two points, it contains the segment joining these points [1, p. 2, and 2. Numbers in square brackets refer to the bibliography at the end]. Minkowski defined certain points of convex sets which he called extreme points [1, pp. 15–16; 3, p. 157]. They are related to certain other points which are here called *extreme points in the sense of distance* to distinguish them from the former, which are called *extreme points in the sense of Minkowski*. A detailed study is made of these two types of extreme points of convex sets in abstract normed linear spaces.

In the first place, it is necessary to distinguish two types of normed linear spaces on the basis of the convexity properties of spherical neighborhoods (§1). A normed linear space such that the segment joining any two points of a spherical neighborhood is interior to the neighborhood except at most for the given points themselves is called a space L^*. All other normed linear spaces are classed together and denoted by L. The study of extreme points is far simpler in spaces L^* than in spaces L, and the results are more complete. An example considered in §10 shows that the property of being a space L^* may depend on the properties of the distance function alone and not on the linearity properties of the space.

In §2 the existence of extreme points is considered. An approximation theorem first proved by Minkowski for euclidean 3-space is extended to spaces L^* and L in §3. Two kinds of convex sets are distinguished in §4 on the basis of the relation of the two kinds of extreme points, and it is shown that the set of extreme points in the sense of Minkowski may be either closed or not closed. In §5 the closed convex hull of a given set is considered, and Minkowski's Approximation Theorem (§3) is extended.

A general theorem on compact sets is established in §6. It is shown that in a complete metric space a set is compact if it is possible to approximate uniformly to it by means of closed compact sets. This theorem and Minkowski's Approximation Theorem (§5) enable us to show in §7 that the closed convex hull of a compact set in a Banach space is compact.

The significance of Minkowski's Approximation Theorem is considered briefly in §8. A series of theorems is given in §9 which establish more precisely the relation between a convex set and its extreme points. Some examples are considered in §10.

The paper may be considered a study in the geometry of abstract space.

1. **Linear spaces and extreme points.** A space which is linear and normed will be designated by L, its elements or points represented by $x, y, \cdots$, and

Received September 10, 1936.

the distance between x and y by $\| x - y \|$. Any spherical neighborhood $\| x - x_0 \| \leqq r$ is convex, since the distance function satisfies the triangle inequality. If the space L has the additional property that no point of the boundary $\| x - x_0 \| = r$ of the sphere $\| x - x_0 \| \leqq r$ is an interior point of a segment joining two other points of the sphere, it will be designated by L^* and called *a space L with non-flat spherical boundaries*. It can be shown at once that no point of the boundary of a sphere is an interior point of a segment joining two interior points of the sphere.

Any euclidean space of n dimensions is a space L^*. A sufficient condition that a space L be a space L^* is that the distance function have the properties of the distance function in Hilbert space.

There are spaces L which are not L^*; for example, the space of continuous functions $f(s)$, $a \leqq s \leqq b$, in which the distance between $f(s)$ and $g(s)$ is max $| f(s) - g(s) |$ on $a \leqq s \leqq b$. For consider the functions $f(s)$ such that $| f(s) | \leqq M$. They form a spherical neighborhood, and two functions $f(s)$, $g(s)$ in it such that $f(a) = g(a) = M$ are on its boundary. But all the functions $\theta f(s) + (1 - \theta) g(s)$, $0 \leqq \theta \leqq 1$, are also contained in the boundary. This space has flat spherical boundaries, i.e., they contain segments.

We shall now define two kinds of extreme points of convex sets.

(1) *An extreme point in the sense of Minkowski* [3, p. 157] *of a convex set C is a point x_M which is not an interior point of any segment joining two points of C.*

(2) *An extreme point in the sense of distance with respect to a point y of a convex set C is a point $x_D(y)$ of C whose distance from y is a maximum.*

As a consequence of the definitions which we have made, we have at once the following theorem.

THEOREM 1.1. *In a space L^* every point of the boundary of a sphere is an extreme point in the sense of Minkowski.*

Notation. In the future C will designate a set which is closed, compact, and convex. The set of extreme points x_M of C will be designated by E_M, and the set of extreme points x_D of C with respect to the points of a set S by $E_D(S)$.

2. Existence of extreme points.

We proceed to establish the existence of the extreme points which we have defined.

THEOREM 2.1. *There exists at least one extreme point $x_D(y)$ of any set C with respect to each point y.*

This theorem is obvious, for the distance from y to a point x of C is a continuous function of x, defined for x on a closed, compact set C.

THEOREM 2.2. *The set of extreme points $E_D(y)$ of C for a fixed y is compact and closed.*

Since $E_D(y)$ is a subset of C, it is compact. That it is closed follows from the remarks made in proof of the last theorem.

Remark. Every set C contains at least one pair of points x_1, x_2 such that $E_D(x_1)$ contains x_2 and $E_D(x_2)$ contains x_1. The segment joining such a pair of points is a diameter of the set.

THEOREM 2.3. *In a space L^* any point x_D of $E_D(S)$ is also a point x_M of E_M.*

The point x_D is related to a point y of S in such a way that the sphere with center y and radius $\| x_D - y \|$ has C in its interior except for certain points on its boundary. From the definition of a space L^* it follows that x_D belongs to E_M.

This theorem together with Theorem 2.1 proves that the set E_M of a set C in a space L^* is not empty.

3. **Minkowski's Approximation Theorem.** In this section we shall establish for spaces L^* a theorem first proved by Minkowski [3, p. 160] for euclidean 3-space, and we shall give its generalization in spaces L.

The convex hull of a finite set of points $x_1, \cdots, x_n$ will be called a polyhedron P_n. Its set of extreme points E_M is composed of the points themselves or a subset of them.

LEMMA 3.1. *Let a set C in L be given and also the parallel set C' at distance $d > 0$. No segment which joins a point of the boundary of C' to an interior point of C' contains a second boundary point of C'.*

The parallel set C' at distance d from C may be defined as the set composed of all points whose distance from C is equal to or less than d (in this case the notation does not imply that C' is compact). Let x_1 be any point of the boundary of C', and x_2 any interior point. Then there exist points y_1 and y_2 of C such that $\| x_1 - y_1 \| = d$, $\| x_2 - y_2 \| < d$. The distance from $\theta x_1 + (1 - \theta)x_2$ to C is not greater than its distance to $\theta y_1 + (1 - \theta)y_2$. But

$$\| \{\theta x_1 + (1 - \theta)x_2\} - \{\theta y_1 + (1 - \theta)y_2\} \| \leqq \theta \| x_1 - y_1 \| + (1 - \theta) \| x_2 - y_2 \| < d$$

for $0 \leqq \theta < 1$. Thus the only point of the segment on the boundary of C' is x_1, and the lemma is established.

THEOREM 3.1. *In a space L^* let any set C with extreme points E_M be given and any $\epsilon > 0$. Then there exist an $N(\epsilon)$ and a sequence of polyhedra $P_1, P_2, \cdots$ with $P_1 \subset P_2 \subset \cdots$, and with all the extreme points in the sense of Minkowski of each one contained in E_M, such that for $n \geqq N(\epsilon)$ the distance from any point of C to P_n is less than ϵ.*

Let $D(x, P_n)$ denote the distance from x to P_n; it is a continuous function of x, for $| D(x, P_n) - D(y, P_n) | \leqq \| x - y \|$.

Let x_1 be any point of E_M, and let x_2 be any point of the set $E_D(x_1)$; at least one exists by Theorem 2.1. Consider $D(x, P_2)$ for x on C. It is a continuous function which is defined on a closed, compact set, and which vanishes on P_2. Unless C is identical with P_2, there is a set of points $X_3 \subset C$, but $X_3 \not\subset P_2$, at which it takes on its maximum value $d_3 > 0$. The set X_3 is closed and compact and lies on the boundary of S_2, the parallel set to P_2 at the distance d_3, and we shall show that it contains at least one point x_3 of E_M. For let x_3 be a point of X_3 whose distance from x_2 is maximum; such a one exists and $\| x_3 - x_2 \| > 0$. Now x_3 is an end point of any segment of C which contains x_3 and a point of C interior to S_2 by Lemma 3.1; all points of C are interior to S_2 except those

of X_3, which are on its boundary. Also x_3 cannot be an interior point of a segment which contains only points of X_3, since these points lie in the interior or on the boundary of the sphere $\| x - x_2 \| \leqq \| x_3 - x_2 \|$. We make use of the fact here that spherical boundaries in a space L^* are non-flat. It follows from the arguments given that $x_3 \epsilon E_M$.

We observe next that $D(x_2, P_1) \geqq D(x_3, P_2)$, for $D(x_3, P_1) \geqq D(x_3, P_2)$, since $P_1 \subset P_2$, and $D(x_2, P_1) \geqq D(x_3, P_1)$, since x_2 is a point of $E_D(x_1)$.

Suppose now that this process has been repeated until n points $x_1, \cdots, x_n$ of E_M have been obtained with $D(x_2, P_1) \geqq D(x_3, P_2) \geqq \cdots \geqq D(x_n, P_{n-1})$. Then it can be repeated to obtain a point x_{n+1} of E_M unless C is identical with P_n. For since C contains points distinct from P_n, $D(x, P_n)$ for x on C takes on its maximum value $d_{n+1} > 0$ at a set of points $X_{n+1} \subset C$, but $X_{n+1} \not\subset P_n$. The set X_{n+1} is closed and compact and lies on the boundary of the parallel set S_n to P_n at the distance d_{n+1}. The same arguments that were used before will show that a point x_{n+1} of X_{n+1} whose distance from x_n is a maximum belongs to E_M.

Also $D(x_n, P_{n-1}) \geqq D(x_{n+1}, P_n)$, for $D(x_{n+1}, P_{n-1}) \geqq D(x_{n+1}, P_n)$ since $P_{n-1} \subset P_n$, and $D(x_n, P_{n-1}) \geqq D(x_{n+1}, P_{n-1})$ because of the definition of x_n.

Thus either there exists a value of n such that C is identical with P_n, or there is an infinite sequence of points $x_1, x_2, \cdots$ of E_M with the corresponding polyhedra $P_1, P_2, \cdots$ and

$$D(x_2, P_1) \geqq D(x_3, P_2) \geqq \cdots \geqq 0. \tag{3.1}$$

Also it is clear that $D(x_n, P_k) \geqq D(x_n, P_{n-1})$ for $k = 1, \cdots, n-1$ (since $P_k \subset P_{n-1}$), from which it follows that

$$\| x_n - x_k \| \geqq D(x_n, P_{n-1}) \qquad (k = 1, \cdots, n-1; n = 2, 3, \cdots). \tag{3.2}$$

The proof will be complete if we can show that the sequence of numbers in (3.1) approaches zero. But if they do not, they have a limit $\delta > 0$ and

$$D(x_n, P_{n-1}) \geqq \delta \tag{3.3}$$

for all values of n. Also the sequence $x_1, x_2, \cdots$ has at least one limit point, since its points belong to the compact set C. Then it is possible to select a sub-sequence which approaches this limit. But no such sub-sequence can approach a limit, because (3.3) and (3.2) show that the distance from any point of it to all the preceding is not less than δ. Thus the assumption has led to a contradiction, and the sequence of numbers in (3.1) approaches zero. The proof of the theorem is complete.

An examination of the proof just given will show that the approximating properties of the polyhedra $P_1, P_2, \cdots$ are in no way dependent on the fact that the points $x_1, x_2, \cdots$ belong to E_M. The hypothesis that C is in L^* was used only to establish the fact that $x_1, x_2, \cdots$ belong to E_M. It will be seen at once that this proof establishes also the following theorem.

THEOREM 3.2. *In a space L let any set C and any $\epsilon > 0$ be given. Then there*

exist an $N(\epsilon)$ and a sequence of polyhedra $P_1, P_2, \cdots$ with $P_1 \subset P_2 \subset \cdots$, all of which are contained in C, such that for $n \geqq N(\epsilon)$ the distance from any point of C to P_n is equal to or less than ϵ.

Remark 3.1. The set C is the limit of the sequence of polyhedra $P_1, P_2, \cdots$ in two senses: (a) take all points x such that $x \in P_n$ for $n \geqq N(x)$ and close this set by adding its limit points; the set obtained is C; (b) $P_n \to C$ according to the definition of Blaschke (see Theorem 9.3 below and [2, p. 60]).

Remark 3.2. By means of the approximating polyhedra, we can set up a denumerable set of points everywhere dense in C; hence C is separable (other proofs are known). An examination of Carathéodory's proof of Blaschke's *Auswahlsatz* shows that it is valid in any separable set; hence the validity of the Weierstrass-Bolzano Cluster Point Theorem for points in a closed convex set implies the validity of the same theorem for closed convex sets [2, pp. 62–66].

Remark 3.3. Let S_1 and S_2 be two closed, compact sets with convex hulls H_1, H_2. If the distance from any point of S_1 to S_2 is at most d, then the distance from any point of H_1 to H_2 is at most d, for the parallel set to H_2 at the distance d obviously includes S_1.

Remark 3.4. The convergence to zero of the sequence of numbers in (3.1) is a necessary condition that the limit of the sequence of polyhedra be a set C, but it is easy to give an example with linear sets to show that it is not sufficient. We consider here a condition that is sufficient. Let a set of points $x(1), x(2), \cdots$ be given; from it form a new sequence $x(k_1) = x(1)$, $x(k_2) = x(2)$, and in general $x(k_n)$ equal to the first element of the original sequence not contained in the convex hull of those that precede it. Let P_n be the convex hull of the first n points of the new sequence. A sufficient condition that the limit of the $P_1, P_2, \cdots$ be a set C, the limit being taken as in Remark 3.1 (a) above, is that the points $x(k_1), x(k_2), \cdots$ belong to some set C. This condition is obviously necessary.

4. **The two kinds of sets** C. As shown by Theorem 2.3, some of the points which are extreme in the sense of distance are also extreme in the sense of Minkowski. On the basis of this relationship we are able to distinguish two kinds of sets C.

THEOREM 4.1. *The set $E_D(S)$ of a set C in a space L is closed if S is compact and closed.*

Let x be any limit point of $E_D(S)$. Then there is a sequence $x_1, x_2, \cdots$, $x_k \in E_D(S)$, whose limit is x. We shall prove that $x \in E_D(S)$.

By hypothesis there is a point $y_k \in S$ such that $x_k \in E_D(y_k)$, $k = 1, 2, \cdots$. Since S is compact and closed by hypothesis, we can select a sub-sequence from the y_k which has a limit y, $y \in S$. Then the limit of the sequence of the corresponding x_k is x, and we may suppose that this selection has been made beforehand so that $y_1, y_2, \cdots$ approaches y.

Suppose that x is not a point of C at maximum distance from y. Then there is a point z, distinct from x, which has this property, and

$$\| z - y \| = \| x - y \| + d, \qquad d > 0.$$

Furthermore

$$\begin{aligned} \| y_n - z \| &\geqq \| y - z \| - \| y_n - y \| \\ &\geqq \| x - y \| + d - \| y_n - y \|. \end{aligned} \tag{4.1}$$

But

$$\| x_n - y_n \| \leqq \| x - y \| + \| x - x_n \| + \| y - y_n \|. \tag{4.2}$$

Since $x_n \to x$, $y_n \to y$, there exists an n sufficiently large, say N, so that

$$\| x - x_N \| < d/4, \; \| y - y_N \| < d/4.$$

Then from (4.1) and (4.2) we have

$$\| y_N - z \| \geqq \| x - y \| + 3d/4,$$

$$\| x_N - y_N \| \leqq \| x - y \| + d/2.$$

These two inequalities contradict the assumption that $x_N \epsilon E_D(y_N)$; hence, there exists no point z of C whose distance from y is greater than that of x. Then $x \epsilon E_D(S)$. The proof is complete.

Definition 4.1. *In a space L^* if there exists a closed compact set S such that the set E_M of a set C is identical with $E_D(S)$, we say C is a set C_S. In all other cases, C is called a set C_X.*

As a result of this definition and Theorem 4.1, we have at once the following theorem.

THEOREM 4.2. *In a space L^* a sufficient condition that the set E_M of C be closed is that C be a set C_S.*

THEOREM 4.3. *In spaces L^* there exist both sets C_S and C_X.*

Any set with a finite number of extreme points E_M, such as a polyhedron, is a set C_S. In particular, every linear set C is a set C_S, the two end points being extreme points in both senses. The sphere $\| x - x_0 \| \leqq r$ is a set C_S also. Its boundary points $\| x - x_0 \| = r$ form E_M (see Theorem 1.1), and S may be taken as the single point x_0.

We shall now give an example of a set C_X. On a circle mark a point x, and let x_1 and x_2 be the end points of a segment which has x for its mid-point and is perpendicular to the plane of the circle. Let x_3 be the point on the circle diametrically opposite x; let x_4 be one of the points on the circle which bisect the arcs joining x_3 and x. In general, let x_k, $k = 5, 6, \cdots$, be the point of the circle which bisects the arc joining x_{k-1} and x. Then the sequence of points $x_1, x_2, \cdots$ has x as its single limit point.

Let P_n denote the convex hull of $x_1, \cdots, x_n$. Then consider the set C of points p such that either (a) $p \epsilon P_n$ for $n \geqq N(p)$, or (b) p is a limit point of the points (a). Then C is closed and compact, lies in a space L^* (see §1), and it is convex also (see Remark 3.4). The points $x_1, x_2, \cdots$ are the extreme points

E_M of C, but E_M is not closed, since their limit point x is not an extreme point. Hence, C is a set C_X, for if it were a set C_S, its extreme points E_M would form a closed set according to Theorem 4.2.

The reader will readily construct plane sets C with points of E_M which are not extreme points in the sense of distance. Thus not every plane set C is a set C_S. In spite of this fact, the following theorem is easily proved.

THEOREM 4.4. *The set E_M of every plane set C is closed.*

Let $p_1, p_2, \cdots, p_n \in E_M \subset C$, have the limit point p. Then the points p_n and p belong to the boundary of C. Let l be the supporting line of C through p. Then C lies entirely on one side of l. If p is the only point of l in C, we see at once that $p \in E_M$, and the theorem is true in this case. If the theorem is false, the only possibility is that p is an interior point of a segment of l which is contained in C. But in this case p would not be a limit point of points of E_M, contrary to hypothesis. Thus the theorem is true in all cases.

5. **On the convex hull of a set.** We shall consider now the definition of the closed convex hull of a set S, and also two methods of constructing it.

Definition 5.1. *The closed convex hull of a set S is the set contained in all closed convex sets which contain S.*

The justification of this definition is the fact that the product of any number of closed convex sets is a closed convex set. Its weakness is that it does not enable us to establish directly many properties of the closed convex hull of S.

Definition 5.2. *The convex hull* [4, p. 359] *of S is the set of points*

$$r_1x_1 + \cdots + r_nx_n,$$

where $x_i \in S$, $r_i \geqq 0$, and $r_1 + \cdots + r_n = 1$.

THEOREM 5.1. *The closed convex hull of S is the closure of the convex hull of S.*

Still another method can be used to construct the closed convex hull of S in the special case that S is compact and lies in L.

Let $\bar{S}$ be the closure of S. Then $\bar{S}$ is also compact [5, p. 89]. Let x be any point of L, and x_1 one of the points of $\bar{S}$ at maximum distance from x (it exists since $\bar{S}$ is compact). Let x_2 be one of the points of $\bar{S}$ at maximum distance from x_1. The convex hull of x_1, x_2 is a polyhedron P_2. We can continue as in the proof of Minkowski's Approximation Theorem (the only difference is that C is replaced by $\bar{S}$) and show that either $\bar{S}$ is contained in some polyhedron P_n, or that there exists a denumerable sequence of polyhedra $P_1 \subset P_2 \subset \cdots$ with all the points E_M of each one contained in $\bar{S}$. Let C be defined as the closure of the set of points x such that $x \in P_n$ for $n \geqq N(x)$. It will be shown later that this set is compact; hence, the notation C is justified.

THEOREM 5.2. *Let C and P_1, P_2, $\cdots$ be the sets whose construction we have just explained. Then the closed convex hull of the compact set S in L is C. Furthermore, given any $\epsilon > 0$, there exists an $N(\epsilon)$ such that $D(x, P_n) < \epsilon$ for $n \geqq N(\epsilon)$ and every $x \in C$. Finally, the set E_M of each polyhedron P_n belongs to $\bar{S}$.*

This theorem may be considered an extension of Minkowski's Approximation

Theorem. Not only can we approximate to the convex hull of S by polyhedra, but we can do so by means of polyhedra whose sets E_M belong to $\bar{S}$.

6. A theorem on compact sets. Let $D(x, G)$ denote the distance from x to the closed compact set G.

THEOREM 6.1. *Given a sequence $G_1, G_2, \cdots$ of closed, compact sets in a metric space, and a set G for which the axiom of completeness is satisfied in addition. If for every $\epsilon > 0$ there exists an $N(\epsilon)$ such that $D(x, G_n) < \epsilon$ for $n \geqq N(\epsilon)$ and $x \epsilon G$, then G is compact.*

Since the sets G_n are compact, they are totally bounded [8, p. 108]. This fact and the hypotheses of the theorem show that G is totally bounded. Then since G is complete by hypothesis, it is compact [8, pp. 107–108].

In the next section this theorem will be applied to sets in a Banach space.

7. The convex hull of a compact set in a Banach space. In this section we shall use the results of the last two sections to establish the following theorem.

THEOREM 7.1. *The convex hull of a compact set S in a Banach space is itself compact.*

Since S is compact, we can use the method of Theorem 5.2 for constructing its closed convex hull C. The polyhedra P_n are closed and compact and satisfy all the other hypotheses for the sets G_n in Theorem 6.1. The set C satisfies all the hypotheses for the set G in the same theorem. Thus Theorem 7.1 follows immediately from Theorems 5.2 and 6.1.

This theorem, although discovered independently by the author, was first proved by Mazur [6].

8. The significance of Minkowski's Approximation Theorem. Let S be a compact set in a linear metric space L. Then it is well known that S is separable, i.e., that there exists a denumerable set of points which is everywhere dense in S. We are therefore led to inquire what additional information is given by Minkowski's Approximation Theorem. The answer is the following (see Theorem 5.2): Let a compact set S in L be given. Then there exists a denumerable set of points $x_1, x_2, \cdots, x_k \epsilon \bar{S}$, such that for each $\epsilon > 0$ there is an $N = N(\epsilon)$ for which $\| x - (r_1x_1 + \cdots + r_Nx_N) \| < \epsilon$, where $r_i \geqq 0$, $r_1 + \cdots + r_N = 1$, and x is any point in the closed convex hull of S. Furthermore, *the set of points $x_1, x_2, \cdots$ in general is not everywhere dense in S.*

The set of points $x_1, x_2, \cdots$ is of course not unique. This fact, for example, indicates clearly the great variety of sets of functions, linear combinations of which can be used to approximate to all the functions of a given set.

9. Theorems on the set E_M. We shall now prove several theorems which will establish more precisely the relation between C and E_M.

THEOREM 9.1. *Let C in L^* be given. Then the closed convex hull of E_M is C.*

Let the closed convex hull of E_M be C'. Then since $E_M \subset C$, it is clear that $C' \subset C$, and the proof will be complete if we can show that $C \subset C'$.

Suppose $C \not\subset C'$. Then $D(x, C')$, which is defined and continuous for x in the closed, compact set C, takes on its maximum value $d > 0$ on a closed compact set $X \subset C$, $X \not\subset C'$. Then the set X lies on the boundary of the parallel set S to C' at the distance d. Let x be a point of X at maximum distance from some point y of C'. Then by the arguments used in the proof of Theorem 3.1, we can show that x is a point of E_M of C. But this is impossible, because E_M is contained entirely in C'. As a result of this contradiction we conclude that $C \subset C'$, and the proof is complete.

Corollary 9.1. *A set C in L^* is determined by any subset E'_M of E_M whose closure contains E_M.*

Theorem 9.2. *Let $x_1, \cdots, x_n$ be points of a set C in L and $r_1, \cdots, r_n$ numbers greater than zero whose sum is 1, and let $x = r_1x_1 + \cdots + r_nx_n$. Then $x \in C$, but if $n > 1$, x is not a point of E_M.*

If $n > 1$, then $0 < r_i < 1$, $i = 1, \cdots, n$, and we can find positive numbers $\epsilon_1, \cdots, \epsilon_n$ such that $0 < r_i \pm \epsilon_i < 1$, and such that $\epsilon_n = \epsilon_1 + \cdots + \epsilon_{n-1}$. Then x is the mid-point of the segment joining the points $(r_1 - \epsilon_1)x_1 + \cdots + (r_{n-1} - \epsilon_{n-1})x_{n-1} + (r_n + \epsilon_n)x_n$ and $(r_1 + \epsilon_1)x_1 + \cdots + (r_{n-1} + \epsilon_{n-1})x_{n-1} + (r_n - \epsilon_n)x_n$, which belong to C, and by definition (1) of §1, x is not an extreme point.

The following theorem is well known [7, §9, p. 200].

Theorem 9.3. *Let $\bar{S}$ be an arbitrary closed bounded set in euclidean N-space and C its closed convex hull. If x is any point of C, then there exist points $x_i \in \bar{S}$ and numbers r_i, $r_i > 0$, $i = 1, \cdots, n$, with $r_1 + \cdots + r_n = 1$, such that $x = r_1x_1 + \cdots + r_nx_n$, and $n \leqq N + 1$.*

Theorem 9.4. *Let $\bar{S}$ be an arbitrary closed bounded set in euclidean N-space, and C its closed convex hull. Then $E_M \subset \bar{S} \subset C$.*

The proof follows at once from the two preceding theorems. Let x be a point of C which is not a point of $\bar{S}$. Then by Theorem 9.3, there are points $x_1, \cdots, x_n$ of $\bar{S}$ such that $x = r_1x_1 + \cdots + r_nx_n$. Also $n > 1$, for otherwise x would be a point of $\bar{S}$ contrary to hypothesis. Then by Theorem 9.2, x is not a point of E_M. It follows that $E_M \subset \bar{S}$, and the proof is complete.

Theorem 9.5. *In euclidean N-space let a set C with extreme points E_M be given. A necessary condition that the closed convex hull C' of $E'_M \subset E_M \subset C$ contain x_M, $x_M \in E_M$ is that x_M be a point or a limit point of E'_M.*

To prove this theorem, we shall suppose that x_M is not a point of $\bar{E}'_M$, and show that it does not belong to C'. Suppose $x_M \in C'$. Then x_M is not an extreme point in the sense of Minkowski of C' because by Theorem 9.4 all such points belong to $\bar{E}'_M$. Then it follows that x_M is not an extreme point in the sense of Minkowski of C, for $C' \subset C$. But this contradicts the hypothesis that $x_M \in E_M \subset C$. Then if $x_M \in C'$, it follows that $x_M \in \bar{E}'_M$, and the proof is complete.

Remark 9.1. It seems likely that Theorem 9.4 is true for sets C in spaces more general than euclidean N-space; hence, we consider the following propositions:

(A) *Let $\bar{S}$ be an arbitrary closed compact set in L, and C its closed compact hull. Then $E_M \subset \bar{S} \subset C$.*

(B) *Let a set C in L with extreme points E_M be given. A necessary condition that the closed convex hull C' of E'_M, $E'_M \subset E_M \subset C$ contain x_M, $x_M \in E_M$, is that x_M be a point of the closure of E'_M.*

An examination of the proof of Theorem 9.5 will show that the following corollary is true.

COROLLARY 9.2. *For every set C in L,* (A) *implies* (B).

Throughout the remainder of this section, we shall treat sets C in L for which (B) holds, no assumption being made about (A). The theorems established will have content because of Theorem 9.5. Among other results, we shall obtain the following restricted converse of Corollary 9.2 (see Corollary 9.5 below): *For every set C in L^*,* (B) *implies* (A).

COROLLARY 9.3. *If C is in L, and if* (B), *a necessary condition that the closed convex hull of $E'_M \subset E_M \subset C$ be C is that the closure of E'_M contain E_M.*

COROLLARY 9.4. *If C is in L^*, and if* (B), *then the closure of the set of points $x_1, x_2, \cdots$ obtained in the proof of Minkowski's Approximation Theorem* 3.1 *contains E_M.*

THEOREM 9.6. *If $C_1, C_2, \cdots$ is a sequence of closed, compact, convex sets in L^* which have the limit C in L^* in the sense of Blaschke, if* (B), *and if $x_M \in E_M \subset C$, then there exists a sequence of extreme points $x'_M \in C_1$, $x''_M \in C_2, \cdots$ whose limit is x_M.*

Let ν_n be the limit inferior of numbers d_n such that the distance from any point of C to C_n, and from any point of C_n to C, is equal to or less than d_n. From the hypothesis that $C_n \to C$ in the sense of Blaschke, it follows that $\nu_n \to 0$ as $n \to \infty$.

Suppose that the theorem is false. Then there exists an $\epsilon > 0$ such that for every $N \geqq 1$ there exists a set C_k with $k \geqq N$ which has no extreme point in the neighborhood $\| x - x_M \| < \epsilon$.

Let S, S_n be the sets of points x of C, C_n such that $\| x - x_M \| \geqq \epsilon/2$, $\| x - x_M \| \geqq \epsilon$ respectively. Let the closed convex hulls of S, S_n be H, H_n. There exists an n, say N_1, such that $\nu_n < \epsilon/2$ for $n \geqq N_1$. Then we can assert that the distance from any point of S_n to S does not exceed ν_n at least for $n \geqq N_1$. Furthermore, by Remark 3.3 the distance from any point of H_n to H does not exceed ν_n for $n \geqq N_1$.

From the hypothesis that proposition (B) holds, it follows that the distance from x_M to H is positive. Call this distance d. Then there exists an n, say $n = N_2$, such that $\nu_n < d/2$ for $n \geqq N_2$. Then for $n \geqq N$, where N is the larger of the integers N_1 and N_2, the distance from any point of H_n to H is less than $d/2$, and therefore the distance from x_M to H_n is greater than $d/2$.

Finally, from the assumption that the theorem is false as stated above it follows that there exists a set C_k, with $k \geqq N$, all of whose extreme points lie in S_k. Then H_k is identical with C_k, because S_k contains only points of C_k, and it contains all of its extreme points E_M (see Theorem 9.1). Thus from the statement made above we see that the distance from x_M to C_k is greater than

$d/2$. But this contradicts the fact that $\nu_k < d/2$. Thus the assumption that the theorem is false has led to a contradiction, and the theorem is established.

Remark 9.2. In the euclidean plane let C_n be the convex hull of the points $(0, 1 + 1/n)$, $(0, -1 - 1/n)$, $(1/n, 0)$, $(-1/n, 0)$. The limit set C is the line segment joining $(0, 1)$ and $(0, -1)$, and these two points are its only extreme points. The point $(0, 0)$ is also a limit point of points x_M of the sets C_n. Thus not every limit point of points x_M of the sets C_n is a point x_M of C. For linear sets, however, every limit point of extreme points is an extreme point of the limit set C.

COROLLARY 9.5. *If S is any compact set in a space L^*, and if* (B), *then the set E_M of the closed convex hull C of S is contained in $\bar{S}$.*

The proof of this corollary follows from Theorems 9.6 and 5.2, for it was shown in the latter that it is possible to construct a sequence of polyhedra whose sets E_M belong to $\bar{S}$, and whose limit in the sense of Blaschke is C.

From this corollary we obtain the following result, stated above.

COROLLARY 9.6. *For every set C in L^*,* (B) *implies* (A).

10. **Some examples.** Consider again the functions $f(s)$ discussed in §1 as an example of a space L which is not L^*. The functions $|f(s)| \leqq M$ form a closed, not compact, convex set. Every function $f(s)$ such that $\max |f(s)| = M$ is an extreme element in the sense of distance, but it can be shown easily that the only extreme elements in the sense of Minkowski are the two functions $f(s) \equiv \pm M$. The closed convex hull of the set E_M is not the given closed convex set (see Theorem 9.1).

Let x be the real number triple (x_1, x_2, x_3), and let $\|x\| = \max |x_i|$, $i = 1, 2, 3$. Then the "spherical neighborhood" $\|x\| \leqq r$ is the cube with vertices at the points $(\pm r, \pm r, \pm r)$; it is a closed, compact, convex set. The space is a space L but not L^*. We thus have an example of a space which is an L^* with the usual euclidean distance function and not an L^* with another one. In both cases, however, the vertices of the cube form its set E_M, and their closed convex hull is the cube itself. Thus the conditions in Theorem 9.1 are sufficient, but not necessary. We see also that the extreme points in the sense of distance of the closed convex hull of a set S need not belong to $\bar{S}$ (see Corollary 9.5).

So far we have characterized an extreme point x_M by the original definition in §1 and by Theorem 9.2. One might expect to characterize x_M at least in euclidean space also by some property of the supporting plane through the point, but this seems impossible. Consider, for example, a plane set C whose boundary contains two intersecting straight line segments and an arc of a circle tangent to one of them. Then we see that there may be either one or many supporting lines through a point x_M which have a single point in common with C, or a single one through x_M, which, however, has a segment in common with C. If a supporting line contains only a single point of C, then that point is a point x_M, however.

Bibliography

1. Bonnesen and Fenchel, *Theorie der konvexen Körper*, Ergebnisse der Mathematik und Ihrer Grenzgebiete, Berlin, vol. 3 (1934), no. 1.
2. Blaschke, *Kreis und Kugel*, Veit und Comp., Leipzig, 1916.
3. Minkowski, *Gesammelte Abhandlungen*, Teubner, Leipzig and Berlin, 1911.
4. Garrett Birkhoff, *Integration of functions with values in a Banach space*, Transactions of the American Mathematical Society, vol. 38 (1935), pp. 357–378.
5. Sierpiński, *General Topology*, The University of Toronto Press, Toronto, 1934.
6. Mazur, *Über die kleinste konvexe Menge, die eine gegebene kompakte Menge enthält*, Studia Mathematica, vol. 2 (1930), pp. 7–9.
7. Carathéodory, *Über den Variabilitätsbereich der Fourier'schen Konstanten von positiven harmonischen Funktionen*, Rendiconti del Circolo Matematico di Palermo, vol. 32 (1911), pp. 193–217.
8. Hausdorff, *Mengenlehre*, Berlin, 1927.

Brown University.

ABELIAN GROUPS WITHOUT ELEMENTS OF FINITE ORDER*

By Reinhold Baer

An abelian group which is written so that its symbols are combined by addition and which has no elements of finite order other than 0[1] may be called completely reducible, if it is a direct sum of groups of rank one. For every group is contained in a completely reducible group of the same rank. There exist furthermore direct irreducible groups of every finite rank and the groups of rank 1 are exactly the subgroups of the additive group of the rational numbers and therefore irreducible.

The structure of a completely reducible group is uniquely determined by the ranks of the differences of certain characteristic subgroups. A survey of the structures of all subgroups of completely reducible groups would involve the solution of the general structure problem, since every group is contained in a completely reducible group. But it is possible to characterize a class of completely reducible subgroups (of completely reducible groups) which are isomorphic with a direct summand of the whole group.

Every property of a completely reducible group which refers to finite subsets or to subgroups of finite rank also holds true for separable groups, i.e., for groups whose finite subsets are contained in completely reducible direct summands. Countable separable groups are completely reducible. But there exist separable groups which are not completely reducible, e.g., vector groups like the additive group of all the sequences of integers. Further criteria for complete reducibility, for separability and for the complete reducibility of separable groups are given. There exists in particular a characterization of the direct summands of finite rank which holds true in every separable group and which is valid in a group of finite rank if, and only if, this group is completely reducible. Furthermore, every direct summand of finite rank of a separable group is completely reducible.

The subsets b' and b'' of the group J are isotype in the group J, if there exists a proper automorphism of J which maps b' upon b''. The classes of isotype elements of a separable group are determined by complete sets of invariants and an enumeration of the characteristic and of the strictly characteristic subgroups is based on this classification of the elements. The existence of characteristic subgroups which are not strictly characteristic and the existence of elements which are not isotype though contained in the same characteristic and in the same strictly characteristic subgroups are closely related phenomena; but neither is a consequence of the other.

Received November 3, 1936; presented to the American Mathematical Society September 3, 1936.

[1] The word "group" is substituted for this longer statement wherever there is no danger of confusion.

A subgroup S of the group J is completely reducible in J, if some direct decompositions of J induce complete reductions of S. Not even all subgroups of rank one of a completely reducible group J are completely reducible in J. But it is possible to give criteria for the complete reducibility of subgroups of finite rank in separable groups. If the subgroup S of J is completely reducible in J, and if the rank of S is finite, then the type of S in J can be described by means of invariants which are derived from the existence of characteristic subgroups of S in characteristic subgroups of J.

All properties and concepts used are invariants and their definitions are based on the concept of multiplicity of an element x in the group J, i.e., the l.c.m. of all the positive integers n such that $x \equiv 0 \bmod nJ$. The calculus with these generalized numbers (in the sense of E. Steinitz) is the basis of the methods applied here.

Previous work in this domain concerns mainly direct sums of a finite number of infinite cyclic groups. They form a rather special, though important, class of completely reducible groups and their theory is embodied in the theory of completely reducible and separable groups as developed here.

Chapter I. Preliminaries

1. **Dependence, rank and rational multipliers.**[2] Let J be an abelian group whose elements are combined by addition and which contains no non-zero element of finite order. If x is an element of J and n an integer, then $nx = 0$ implies that $n = 0$ or $x = 0$. An element x of J is therefore *dependent* on the subset S of J, if there exists a positive integer n such that nx is contained in the subgroup of J which is generated by the elements in S. A subset of J is *dependent*, if at least one of its elements depends on the others. A subset of J is therefore independent if, and only if, all its finite subsets are independent, and the finite subset $b_1, \cdots, b_k$ is independent if, and only if, $\sum_{i=1}^{k} c_i b_i = 0$ implies that all the integral coefficients c_i are 0.

The *rank* of J is the smallest (finite or infinite) number $r(J)$ such that there exists a subset of J on which every element of J is dependent and which contains $r(J)$ elements. There always exist greatest independent subsets of J. If G is a greatest independent subset of J, then every element of J is dependent on G and G contains exactly $r(J)$ elements.

The subset C of the group J is *closed* (in J), if C contains every element of J which is dependent on C. Closed subsets are subgroups and a subgroup S of J is closed in J if, and only if, the class group J/S does not contain elements $\neq 0$ of finite order.

Since the intersection of any number of closed subgroups is also a closed subgroup, there always exists a smallest closed subgroup which contains a given

[2] For proofs of the facts mentioned in this section, see R. Baer, *The subgroup of the elements of finite order of an abelian group*, Ann. of Math., vol. 37 (1936), pp. 766–781.

set S: *the closed subgroup, generated by* S. It contains exactly those elements which are dependent on S.

In groups of rank one any pair of elements is dependent. The elements $x \neq 0$ and $y \neq 0$ are dependent if, and only if, there exist integers n and m such that

$$nx = my \neq 0.$$

The ratio of n and m is uniquely determined by the dependent elements x and y; and to given x, n, m there exists at most one solution y of the equation $nx = my$. If the equation $nx = my$ has a solution y in J, the notation $y = \frac{n}{m}x$ may be used.

By this definition a multiplication of the elements of J with rational numbers has been introduced. Since it is possible that rb' exists in J, but rb'' does not, the rational numbers are not operators in the usual sense. These rational multipliers satisfy the following rules.

Suppose that r and s are rational numbers, x and y elements in J.

If rx and ry exist in J, then $r(x \pm y)$ exists in J and satisfies

$$r(x \pm y) = rx \pm ry.$$

If rx and sx exist in J, then $(r \pm s)x$ exists in J and satisfies

$$(r \pm s)x = rx \pm sx.$$

If rx and $s(rx)$ exist in J, then $(sr)x$ exists in J and satisfies $(sr)x = s(rx)$.

If rx and sx exist in J, and if the denominators of r and s are relatively prime, then $(rs)x$ exists in J and satisfies

$$(rs)x = r(sx) = s(rx).$$

A proof of the last formula may be added. If $r = mn^{-1}$, $s = hk^{-1}$ and n and k are relatively prime, then there exist integers k', n' such that $nn' + kk' = 1$ and therefore

$$mhx = mhkk'x + mhnn'x = kn(hk'rx + mn'sx).$$

It is important to note that the converses of these rules are not true.

These formulas show that the closed subgroup of J which is generated by the element $x \neq 0$ is simply isomorphic with the additive group of those rational numbers which are multipliers of x. Groups of rank one are therefore exactly the subgroups of the additive group of all the rational numbers. Hence they may be called *rational groups*.

The group J is *completely reducible*, if J is a direct sum of rational groups. If J is the direct sum of the rational groups J_v and $b_v \neq 0$ an element of J_v, the elements b_v form a *basis* of J. The subset B of J is therefore a basis of J if, and only if, B is a greatest independent subset of J and an equation

$$nx = \sum_{i=1}^{k} c_i b_i$$

where n is a positive integer, $x \neq 0$ an element in J, the c_i are integers and the b_i different elements in B, implies that every $(c_i n^{-1})b_i$ exists in J.

J is *complete*, if $J = nJ$ for every positive integer n. Complete subgroups are direct summands. Complete groups are direct sums of groups which are isomorphic with the additive group of all the rational numbers. Every greatest independent subset of a complete group is a basis. Every group is contained in an essentially uniquely determined smallest complete group.

2. **Multiplicities and derived invariants.** In the following a certain generalization of the multiplicative set of the positive integers will be needed.[3] The principal concepts and properties of these generalized numbers will therefore be enumerated.

If, for every prime number p, $v(p)$ is either a non-negative integer or the symbol ∞, then

$$v = \prod_p p^{v(p)}$$

is such a generalized number v and v is uniquely determined by its p-values $v(p)$.

If S is any set of (ordinary or generalized) numbers, the product of the numbers in S is the number whose p-value is the sum of the p-values of the numbers in S. Here the sum of an infinity of positive integers is ∞ and the sum of ∞ and anything is ∞. The p-value of the g.c.d. (l.c.m.) of the numbers in the set S is the minimum (maximum) of the p-values of the numbers in S.

v is a divisor of w, w a multiple of v, i.e., $v \mid w$, if there exists a solution x of the equation $w = xv$. $v \mid w$ if, and only if, $v(p) \leqq w(p)$ for every prime number p (where $v(p)$ is the p-value of v and $w(p)$ the p-value of w). If $v \mid w$, there may exist many solutions of the equation $w = xv$. The g.c.d. $w :_s v$ and the l.c.m. $w :_g v$ of all the solutions x of $w = xv$ are also solutions of this equation, the smallest and the greatest, respectively. If $v \mid w$, the p-value of

$w :_s v$ is	$w :_g v$ is	
$w(p) - v(p)$	$w(p) - v(p)$	if $w(p)$ is finite
∞	∞	if $w(p)$ is infinite, $v(p)$ finite
∞	0	if $v(p)$ is infinite.

The infinite part of the number v is $v_\infty = v :_g v$ and the finite part is $v_f = v :_s v_\infty$. The symbols v_∞ and v_f are relatively prime, the p-values of v_∞ either 0 or ∞ and the p-values of v_f are all finite. We have $v = v_f v_\infty$.

1 is the g.c.d. of all the numbers. Every number is the l.c.m. of ordinary integers and the l.c.m. of all numbers is the number without finite p-values. It will be convenient to add to these numbers a symbol ∞ which is different from all numbers and a multiple of every number.

The numbers v and w have *the same genus*, i.e., $|v| = |w|$, if there exist ordinary positive integers m and n such that $mv = nw$. Hence $|v| = |w|$ if, and only if, $v_\infty = w_\infty$ and for almost every p the p-value of v is the same as the p-value of w.

[3] E. Steinitz has used this generalization for enumeration of the finite fields.

If $|v_i| = |w_i|$, then $|v_1 v_2| = |w_1 w_2|$ and

$$|v_1 \dot{:}_s v_2| = |w_1 \dot{:}_s w_2|, \qquad |v_1 \dot{:}_g v_2| = |w_1 \dot{:}_g w_2|.$$

Product, g.c.d. and l.c.m. of a finite number of genera are therefore uniquely determined genera.

If the three genera $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ satisfy $\mathbf{a} = \mathbf{bc}$, then $\mathbf{b}$ is a divisor of $\mathbf{a}$ and $\mathbf{a}$ a multiple of $\mathbf{b}$, i.e. $\mathbf{b} \leqq \mathbf{a}$. If $\mathbf{b} \leqq \mathbf{a}$ and $\mathbf{a} \leqq \mathbf{b}$, then $\mathbf{a} = \mathbf{b}$. If $\mathbf{a} \leqq \mathbf{b}$ and $\mathbf{a} \neq \mathbf{b}$, then $\mathbf{a} < \mathbf{b}$. The genus $\mathbf{a}$ is a divisor of the genus $\mathbf{b}$ if, and only if, there exist numbers of genus $\mathbf{a}$ which are divisors of certain numbers of genus $\mathbf{b}$.

DEFINITION 2.1. *The multiplicity* $m(x) = m(x < J)$ *of the element* x *in the group* J *is* ∞, *if* $x = 0$, *and is the* l.c.m. *of all the positive integers* n *such that* $x \equiv 0 \bmod nJ$ (*i.e., such that* $n^{-1}x$ *exists in* J), *if* $x \neq 0$.

If $x \neq 0$, *the genus* $|x| = |x < J|$ *of* x *in* J *is the genus of the multiplicity of* x *in* J.[4]

If x is an element of the subgroup S of J, then $m(x < S)$ is a divisor of $m(x < J)$ and, if $x \neq 0$, then $|x < S|$ is a divisor of $|x < J|$. If S is a closed subgroup of J, the elements of S have the same multiplicity and the same genus in S and in J.

If S is a (closed) subgroup of J, x an element of J, the multiplicity $m(x < J/S) = m(S + x < J/S)$ of x mod S is a multiple of $m(x < J)$ and, if $x \not\equiv 0 \bmod S$, the genus $|x < J/S| = |S + x < J/S|$ of x mod S is a multiple of its genus in J.

If n is a positive integer, then $m(\pm nx) = nm(x)$ and $n^{-1}x$ exists in J if, and only if, n is a divisor of $m(x < J)$. If r is a positive rational number such that rx exists in J, then $m(rx) = rm(x)$.

(2.2) (a) *If* $x = \sum_{i=1}^{k} x_i$, *then* $m(x)$ *is a multiple of the* g.c.d. *of the* $m(x_i)$ *and, if* x, x_i *are all* $\neq 0$, *then* $|x|$ *is a multiple of the* g.c.d. *of the* $|x_i|$.

(b) *If* $x = \sum_{i=1}^{k} x_i$, *and if the elements* $x_i \neq 0$ *are contained in different components of a direct decomposition of the group* J, *then* $m(x)$ *is the* g.c.d. *of the* $m(x_i)$ *and* $|x|$ *the* g.c.d. *of the* $|x_i|$.

(c) *The elements* $x \neq 0$ *and* $y \neq 0$ *generate isomorphic closed subgroups of the group* J *if, and only if,* $|x| = |y|$.

(d) *The closed subgroup of* J, *generated by the element* $x \neq 0$, *is isomorphic with a subgroup of the closed subgroup of* J, *generated by the element* $y \neq 0$, *if and only if,* $|x| \leqq |y|$.

Proof. (a) and (b) are consequences of the corresponding facts for the p-values of the multiplicities. If x and y are contained in the same group of rank one, there exist integers n and u such that $nx = uy \neq 0$ and this implies

[4] The assumption that there do not exist elements $\neq 0$ of finite order in J is not needed for this definition. If J is an abelian group without elements of infinite order, the multiplicity is essentially Prüfer's "Höhe".

the necessity of the conditions in (c) and (d). If finally the condition of

(c)	(d)

is satisfied, there exist integers $n \neq 0$, $u \neq 0$ such that

$m(nx) = m(uy)$.	$m(nx) \mid m(uy)$.

If therefore $r \neq 0$ is a rational number, then

$r(uy)$ exists if, and only if, $r(nx)$ exists.	$r(uy)$ exists, if $r(nx)$ exists.

Hence in mapping $r(nx)$ upon $r(uy)$ a required isomorphism is defined.

By (c) all the elements $\neq 0$ of a rational group have the same genus. The genus of the elements $\neq 0$ of the rational group R is called the *genus* $|R|$ of R.

(c) and (d) imply that the rational groups R and R' are isomorphic, if R is isomorphic with a subgroup of R' and R' with a subgroup of R.

DEFINITION 2.3. *If $f(x)$ is a property, then $(J, f(x))$ is the (not necessarily closed) subgroup of J, generated by the elements x of J which satisfy $f(x)$.*

$f(x)$ is an additive property, if[5]

$$\left(\sum_v J_v, f(x)\right) = \sum_v (J_v, f(x)).$$

The following properties $f(x)$ will be used: $g \mid m(x < J)$; the order of x mod the subgroup S of J is a finite divisor of g; $\mathbf{s} \leqq |x < J|$; $\mathbf{s} < |x < J|$; $|x < J| \nless \mathbf{s}$; $|x < J| \nleqq \mathbf{s}$. Here g is an (ordinary or generalized) number and $\mathbf{s}$ a genus.

All these properties (except the second) are additive. If $f(x)$ is one of the three first properties, every element of $(J, f(x))$ has the property $f(x)$.

Since $(J, g/m(x < J))$ is the intersection of the groups nJ for finite divisors n of g, and since its structure depends only on g and on the structure of J, it may be denoted by gJ.

The closed subgroup of J, generated by gJ, is exactly $(J, |g| \leqq |x < J|)$ and the groups $(J, \mathbf{s} \leqq |x < J|)$ are therefore closed subgroups of J.

If $f(x, g, S)$ is the second property, and if $g \mid m(x < J)$ for every x in S, then $(J, f(x, g, S))$ is the join of the subgroups $n^{-1}S$ of J for finite divisors n of g. Since under this assumption the structure of $(J, f(x, g, S))$ depends only on g and on the structure of S, and since every group is contained in a complete group, this property may be used for defining $g^{-1}S$. If $g_\infty = 1$, then $g^{-1}(gJ) = J$; if furthermore $J = (J, |g| \leqq |x|)$, then $g(g^{-1}J) = J$. Thus under these assumptions the structures of J and gJ and the structures of J and $g^{-1}J$ determine each other. Note that for a rational group R and a prime number p either $p^\infty R = R$ or $p^\infty R = 0$.

Finally, $m(x < g^{-1}J) = gm(x < J)$ and $gm(x < gJ) = m(x < J)$.

(2.4) (a) $(J, \mathbf{s} < |x|) \leqq (J, \mathbf{s} \leqq |x|)$, $(J, |x| \nleqq \mathbf{s}) \leqq (J, |x| \nless \mathbf{s})$,

$(J, \mathbf{s} < |x|) \leqq (J, |x| \nleqq \mathbf{s})$, $(J, \mathbf{s} \leqq |x|) \leqq (J, |x| \nless \mathbf{s})$.

[5] $\sum_v J_v$ is here and in the future the *direct* sum of the groups J_v.

(b) Every class of $J(\mathbf{s})^* = (J, \mathbf{s} \leqq |x|)/(J, \mathbf{s} < |x|)$ is contained in exactly one class of $J(\mathbf{s})^{**} = (J, |x| \not< \mathbf{s})/(J, |x| \not\leqq \mathbf{s})$. Thus a *homomorphism* of $J(\mathbf{s})^*$ upon the whole group $J(\mathbf{s})^{**}$ is defined and *this homomorphism is an isomorphism if, and only if,*

(b*) $(J, \mathbf{s} < |x|)$ *is exactly the intersection of* $(J, \mathbf{s} \leqq |x|)$ *and of* $(J, |x| \not\leqq \mathbf{s})$.

These statements are consequences of the definitions and of the following fact. If y is an element of $(J, |x| \not< \mathbf{s})$, then y is the sum of elements y_i and of elements z_j such that $\mathbf{s} \leqq |y_i|$ and $\mathbf{s} \not\leqq |z_j| \not< \mathbf{s}$. Since the y_i are therefore contained in $(J, \mathbf{s} \leqq |x|)$ and the z_j in $(J, |x| \not\leqq \mathbf{s})$, this implies that every element of $(J, |x| \not< \mathbf{s})$ is mod $(J, |x| \not\leqq \mathbf{s})$ congruent to an element of $(J, \mathbf{s} \leqq |x|)$.

DEFINITION 2.5. *The direct decomposition* $J = \sum_v J_v$ *is a partial reduction of* J, *if for every* v *all the elements* $\neq 0$ *of* J_v *have the same genus in* J_v *(and in* J*).*

Every complete reduction of the group J is by (2.2c) also a partial reduction. If $J = \sum_v J_v$ is a partial reduction of the group J, let $J(\mathbf{t})$ be the sum of all those J_v whose elements $\neq 0$ have the genus $\mathbf{t}$. Then $J = \sum_{\mathbf{t}} J(\mathbf{t})$ is also a partial reduction of J, since every element $\neq 0$ in $J(\mathbf{t})$ has the genus $\mathbf{t}$. Since, therefore, elements $\neq 0$ in different components $J(\mathbf{t})$ have different genera, this decomposition is a *smallest partial reduction of* J.

(2.6) *If* $J = \sum_{\mathbf{t}} J(\mathbf{t})$ *is a smallest partial reduction of* J, *then*

$$(J, \mathbf{t} \leqq |x|) = \sum_{\mathbf{t} \leqq \mathbf{s}} J(\mathbf{s}), \qquad (J, \mathbf{t} < |x|) = \sum_{\mathbf{t} < \mathbf{s}} J(\mathbf{s}),$$

$$(J, |x| \not< \mathbf{t}) = \sum_{\mathbf{s} \not< \mathbf{t}} J(\mathbf{s}), \qquad (J, |x| \not\leqq \mathbf{t}) = \sum_{\mathbf{s} \not\leqq \mathbf{t}} J(\mathbf{s})$$

and $J(\mathbf{t})$ *represents exactly the classes of* $J(\mathbf{t})^*$ *and the classes of* $J(\mathbf{t})^{**}$.

These are consequences of (2.2) and (2.4).

(2.7) *Suppose that every finite subset of the group* J *is contained in a partially reducible direct summand of* J.

(a) $(J, \mathbf{t} < |x|)$, $(J, |x| \not< \mathbf{t})$ *and* $(J, |x| \not\leqq \mathbf{t})$ *are closed subgroups of* J.

(b) $(J, \mathbf{t} < |x|)$ *is for every genus* $\mathbf{t}$ *the intersection of* $(J, \mathbf{t} \leqq |x|)$ *and* $(J, |x| \not\leqq \mathbf{t})$.

(c) *All the elements* $\neq 0$ *of* $J(\mathbf{t})^*$ *and of* $J(\mathbf{t})^{**}$ *have the genus* $\mathbf{t}$.

(d) *To every element* $b \neq 0$ *of* J *there exists at least one and at most a finite number of genera* $\mathbf{t}$ *such that*

$$b \equiv 0 \bmod (J, |x| \not< \mathbf{t}), \qquad b \not\equiv 0 \bmod (J, |x| \not\leqq \mathbf{t}).$$

Proof. If b is any element of J, there exists a partially reducible direct summand D of J which contains b. If $f(x)$ is one of the three discussed properties, then $(D, f(x))$ is a direct summand of $(J, f(x))$ and by (2.6) a direct summand of D and now (a)–(c) are consequences of (2.6) and (2.4). If $b \neq 0$ and $D = \sum_{\mathbf{t}} D(\mathbf{t})$ is a smallest partial reduction of J, then $b = \sum_{i=1}^{k} b_i$, $b_i \neq 0$, b_i in $D(\mathbf{t}_i)$,

$\mathbf{t}_i \neq \mathbf{t}_j$ for $i \neq j$. The genera $\mathbf{t}$ such that $b \equiv 0 \bmod (J, |x| \not< \mathbf{t})$, $b \not\equiv 0 \bmod (J, |x| \not\leqq \mathbf{t})$ are exactly the genera $\mathbf{t}_i$ such that $\mathbf{t}_j \not< \mathbf{t}_i$ for every $j \neq i$, and this proves (d).

THEOREM 2.8. (a) *Suppose that J_i is a direct sum of rational groups of genus $\mathbf{t}_i$. Then J_1 and J_2 are isomorphic if, and only if,*

$$\mathbf{t}_1 = \mathbf{t}_2, \qquad r(J_1) = r(J_2).$$

(b) *Suppose that J_i is partially reducible. Then J_1 and J_2 are isomorphic if, and only if, for every genus $\mathbf{t}$, $J_1(\mathbf{t})^*$ and $J_2(\mathbf{t})^*$ are isomorphic.*

(c) *Suppose that J_i is completely reducible. Then J_1 and J_2 are isomorphic if, and only if, for every genus $\mathbf{t}$*

$$r(J_1(\mathbf{t})^*) = r(J_2(\mathbf{t})^*).$$

Proof. If J_i is a direct sum of rational groups of genus $\mathbf{t}_i$, the number of summands is $r(J_i)$ and $(J_i, \mathbf{t}_i < |x|) = 0$, $J_i = (J_i, \mathbf{t}_i \leqq |x|)$. This proves (a). (b) is a consequence of (2.6) and (c) is a consequence of (a), (b), (2.6).

COROLLARY 2.9. *Any two decompositions of a group J into rational direct summands are isomorphic.*

This is true since there appear exactly $r(J(\mathbf{t}))$ summands of genus $\mathbf{t}$ in such a decomposition.

Chapter II. Direct summands of finite rank

3. **Direct summands and complete reducibility of groups of finite rank.** If the group J is the direct sum of its subgroups S and T, i.e., $J = S + T$, then

$$(J, \mathbf{s} \leqq |x|) = (S, \mathbf{s} \leqq |x|) + (T, \mathbf{s} \leqq |x|),$$
$$(J, \mathbf{s} < |x|) = (S, \mathbf{s} < |x|) + (T, \mathbf{s} < |x|).$$

The subgroup of $J(\mathbf{s})^*$ which contains elements of the direct summand S of J is therefore a direct summand of $J(\mathbf{s})^*$. Hence a direct summand S of the group J satisfies the following conditions.

CONDITION 3.1. *The classes of $J(\mathbf{s})^*$ which contain elements of S form a closed subgroup of $J(\mathbf{s})^*$.*[6]

CONDITION 3.2. *$(S, \mathbf{s} < |x|)$ is the intersection of S and $(J, \mathbf{s} < |x|)$.*

The above formulas imply:

(3.3) The subgroup S of the direct summand D of the group J satisfies the conditions 3.1 and 3.2 in J if, and only if, S satisfies these conditions in D.

[6] It may happen that $(J, \mathbf{s} < |x|)$ is not a closed subgroup of J and that consequently $J(\mathbf{s})^*$ contains elements $\neq 0$ of finite order. Then the "closed" subgroups of $J(\mathbf{s})^*$ are not defined and therefore the following proposition may be substituted for condition 3.1, if $(J, \mathbf{s} < |x|)$ is not closed: *a congruence $nz \equiv s \bmod (J, \mathbf{s} < |x|)$, where n is an ordinary positive integer and s an element of S whose genus in J is $\mathbf{s}$, has a solution z in J if, and only if, there exists a solution z in S.*

Note that this proposition and condition 3.1 are equivalent, if $(J, \mathbf{s} < |x|)$ is a closed subgroup of J.

THEOREM 3.4. *The group J of finite rank is completely reducible if, and only if, every subgroup S of J which satisfies the conditions* 3.1 *and* 3.2 *is a direct summand of J.*

Proof. A. Suppose that J is a completely reducible group of finite rank and that the subgroup S of J satisfies the conditions 3.1 and 3.2.

Case 1. J is a direct sum of a finite number of isomorphic groups (of rank 1) and $r(S) = 1$.

By condition 3.1 S is a closed subgroup of J. Thus if $r(J) = 1$, then $J = S$. Suppose now that $r(J) = 2$. Then $J = J' + J''$ and either S is one of these direct summands or there exist elements $b' \neq 0$ in J', $b'' \neq 0$ in J'' and relatively prime positive integers n', n'' such that $n'b' + n''b''$ is an element $\neq 0$ of S and $m(b') = m(b'') = m(n'b' + n''b'')$. There exist therefore integers k', k'' such that $k'n' - k''n'' = 1$. Then the elements

$$c' = n'b' + n''b'', \qquad c'' = k''b' + k'b''$$

form a basis of J, since $m(c') = m(c'') = m(b') = m(b'')$, and since the subgroup of J, generated by c' and c'', contains b' and b''. Thus S is a direct summand of J.

If, finally, $J = \sum_{i=1}^{r(J)} J_i$, where the J_i are isomorphic rational groups, $r(J)$ any finite number, and $s \neq 0$ any element in S, then $s = \sum_{i=1}^{r(J)} s_i$ with s_i in J_i. Then $s' = \sum_{i=2}^{r(J)} s_i$ generates a closed subgroup S' of $J' = \sum_{i=2}^{r(J)} J_i$ and, by complete induction, S' is a direct summand of J'. Thus S is contained in a direct summand of rank 2 and is therefore, as proved above, a direct summand of J.

Case 2. J is a completely reducible group of finite rank and $r(S) = 1$.

Let $\mathbf{s}$ be the genus of S and let S^* be the subgroup of $J(\mathbf{s})^*$ whose classes contain elements of S. Then every element of S is contained in one class of S^* and every class of S^* contains by condition 3.2 exactly one element of S. S^* is by condition 3.1 a closed subgroup of rank one of $J(\mathbf{s})^*$. Since J is completely reducible and of finite rank, and since therefore, by (2.6), $J(\mathbf{s})^*$ is a direct sum of a finite number of isomorphic rational groups, it follows from case 1 that S^* is a direct summand of $J(\mathbf{s})^*$, i.e., $J(\mathbf{s})^* = S^* + T^*$. If T is the subgroup of $(J, \mathbf{s} \leqq |x|)$ which contains $(J, \mathbf{s} < |x|)$ and satisfies $T^* = T/(J, \mathbf{s} < |x|)$, then $(J, \mathbf{s} \leqq |x|) = S + T$, since S represents exactly S^*. Since, by (2.6), $(J, \mathbf{s} \leqq |x|)$ is a direct summand of J, this implies that S is a direct summand of J.

Case 3. J is a completely reducible group of finite rank.

There exist in S elements $w \neq 0$ such that $|w < S| \not< |s < S|$ for every element $s \neq 0$ in S.

For if $b \neq 0$ is an element in S whose genus in S is $\mathbf{s}$, and if $\mathbf{s} < \mathbf{t}$, then $(S, \mathbf{t} \leqq |x|) < (S, \mathbf{s} \leqq |x|)$ and therefore, since both these subgroups are closed subgroups of S, $r((S, \mathbf{t} \leqq |x|)) < r((S, \mathbf{s} \leqq |x|))$.

Let now w be an element $\neq 0$ in S such that $|w < S| \not< |s < S|$ for every

$s \neq 0$ in S, let W be the closed subgroup of S, generated by w, and $\mathbf{s}$ the genus of W. Since $(S, \mathbf{s} < |x|) = 0$, and since S satisfies the conditions 3.1 and 3.2, W satisfies the conditions 3.1 and 3.2 in J and is therefore as group of rank one a direct summand of J, i.e., $J = W + J'$.

Then $S = W + S'$ where S' is the intersection of S and J', since $W \leqq S$. The subgroup S' of J' satisfies therefore the conditions 3.1 and 3.2 in J', and since $r(J') = r(J) - 1$, it can be assumed that S' is a direct summand of J'. Thus it has been proved by complete induction with regard to $r(J)$ that S is a direct summand of J.

B. Suppose now that J is a group of finite rank such that every subgroup of J which satisfies the conditions 3.1 and 3.2 in J is a direct summand of J.

Since J is a group of finite rank, there exists (as proved in A., case 3) an element $w \neq 0$ in J such that $|w < J| \nless |b < J|$ for every element $b \neq 0$ in J. The closed subgroup W of J, generated by w, satisfies the conditions 3.1 and 3.2 and is therefore a direct summand of J, i.e., $J = W + J'$. Every subgroup S of J' which satisfies the conditions 3.1 and 3.2 in J' by (3.3) also satisfies them in J and is therefore a direct summand of J and of J'. Since $r(J') = r(J) - 1$, it can therefore be assumed that J' is completely reducible. Hence the complete reducibility of J has been proved by complete induction with regard to $r(J)$.

Corollary 3.5. *If J is a completely reducible group of finite rank, every direct summand of J is completely reducible.*

For if the subgroup S of the direct summand D of J satisfies the conditions 3.1 and 3.2 in D, it follows from (3.3) and Theorem 3.4 that S is a direct summand of J and therefore a direct summand of D, i.e., D is by Theorem 3.4 completely reducible.

Corollary 3.6. *Suppose that all the elements $\neq 0$ of the group J have the same genus $\mathbf{s}$ in J.*

(a) *If J is complete, then J is completely reducible and every closed subgroup of J is a direct summand of J.*

(b) *If J is not complete, every closed subgroup of J is a direct summand of J if, and only if, J is a completely reducible group of finite rank.*

Proof. (a) is a consequence of the fact that every closed subgroup of a complete group is complete and that complete subgroups are direct summands, complete groups direct sums of rational groups (§1).

Since $J = (J, \mathbf{s} \leqq |x|)$, $0 = (J, \mathbf{s} < |x|)$, the closed subgroups of J satisfy the conditions 3.1 and 3.2. Every closed subgroup of J is therefore by Theorem 3.4 a direct summand of J, if J is a completely reducible group of finite rank. Suppose now that J is not complete and that every closed subgroup of J is a direct summand of J. If $r(J)$ is infinite, then J contains an independent subset $b_1, b_2, \cdots, b_i, \cdots$. The closed subgroup J' of J, generated by these elements, is a direct summand of J. Since every subset of the sequence of the b_i generates a closed subgroup of J which is a direct summand of J and of J', the elements b_i form a basis of J'. Since J is not complete, there exists a prime number p

such that $pJ < J$. The closed subgroup of J and of J' which is generated by the elements $b_{i-1} - pb_i$ is a direct summand J'' of J and of J'. J'' does not contain b_1. But $\bar{J} = J'/J''$ satisfies $p\bar{J} = \bar{J}$. This is a contradiction, since the p-value of $m(b < J)$ is finite for every $b \neq 0$ in J. The rank of J is therefore finite. Since every closed subgroup of J is a direct summand of J, Theorem 3.4 implies the complete reducibility of J.

THEOREM 3.7. *If there exists a (generalized) number g and a finite subset F of the group J such that $g \mid m(f < J)$ for every f in F and such that J is generated by the elements $n^{-1}f$ with f in F and n any positive integer, dividing g, then J is a direct sum of a finite number of rational groups of genus $\mid g \mid$.*

Proof. Denote by $f_1, \cdots, f_k$ the elements of F. Let J' be a direct sum of k rational groups of genus $\mid g \mid$ and $f'_1, \cdots, f'_k$ a basis of J' such that $g = m(f'_i < J')$. There exists a homomorphism α of J' upon the whole group J such that $f'_i\alpha = f_i$. If W' is the subgroup of all the elements of J' which are mapped upon 0 by α, then J'/W' and J are isomorphic. W' is therefore a closed subgroup of J'. W' is by Corollary 3.6 a direct summand of J' and J'/W' is therefore by Corollary 3.5 completely reducible. Hence J is completely reducible and J is by Corollary 2.9 a direct sum of rational groups of genus $\mid g \mid$.

COROLLARY 3.8. *If there exists a finite subset F of J such that all the elements of F have the same genus* $\mathbf{s}$ *in J, and such that J is generated by the rational multiples of the elements in F, then J is a direct sum of a finite number of rational groups of genus* $\mathbf{s}$.

For if g is the g.c.d. of the numbers $m(f < J)$ with f in F, then $m(f < J):g = n(f)$ is an ordinary positive integer and the elements $f' = n(f)^{-1}f$ for f in F form a subset F' of J such that g, F' satisfy the conditions of Theorem 3.7.

COROLLARY 3.9.[7] *Suppose that all the elements $\neq 0$ of the group J of finite rank have the same genus in J. Then J is completely reducible if, and only if, J/B is finite for every subgroup B which is generated by the rational multiples of the elements of a given greatest independent subset of J.*

Proof. Suppose first that J is a direct sum of a finite number of isomorphic rational groups, i.e., $J = \sum_{i=1}^{r(J)} J_i$, and that the subgroup B of J is generated by the rational multiples of the elements of a greatest independent subset G of J. Then every element $b \neq 0$ of B has the same genus $\mathbf{s}$ in B and in J. If therefore J'_i is the intersection of B and J_i, then J_i/J'_i is a finite cyclic group. The direct sum J' of the groups J'_i is a subgroup of B and J/J' is a finite group, since $r(J)$ is finite. Since J/B and $(J/J')/(B/J')$ are isomorphic, J/B is also a finite group.

Suppose now that G is a greatest independent subset of J, that B is the subgroup of J generated by the rational multiples of the elements in G, and that J/B is finite. Let F be a set of elements, containing G and a complete set of representatives of the classes of J/B. Then F is finite and satisfies the assumptions of Corollary 3.8 and J is therefore completely reducible.

[7] A particular case of this proposition has been communicated to the author by Dr. G. Szekeres.

4. Direct summands of finite rank and separable groups.

DEFINITION 4.1. *The group J is separable, if every finite subset of J is contained in a completely reducible direct summand of J.*

Note that the assumptions of (2.7) are satisfied for separable groups.

THEOREM 4.2. *J is a separable group if, and only if,*

(a) *every finite subset of J is contained in a direct summand of finite rank;*

(b) *every subgroup of finite rank which satisfies the conditions* 3.1 *and* 3.2 *is a direct summand of J.*

Proof. If H is the direct sum of the rational groups H_v, then every element of H is contained in the direct sum of a finite number of the groups H_v. Every finite subset and every subgroup of finite rank of a separable group is therefore contained in a completely reducible direct summand of finite rank. Hence every separable group satisfies the conditions (a) and (b) by Theorem 3.4.

If the group J satisfies the conditions (a) and (b), every direct summand of finite rank is by (3.3) and Theorem 3.4 completely reducible and J is therefore by (a) separable.

COROLLARY 4.3. *If J is a separable group, every finite subset of J is contained in a completely reducible direct summand of finite rank and every direct summand of finite rank is completely reducible.*[8]

COROLLARY 4.4. *If all the elements $\neq 0$ of J have the same genus in J, then J is separable if, and only if, every closed subgroup of finite rank in J is a direct summand of J.*

This is a consequence of Theorem 4.2 and Corollary 3.6, since any n elements of J generate a closed subgroup of a rank $\leqq n$.

COROLLARY 4.5. *Denote by C the greatest complete subgroup of the group J. Then every closed subgroup of finite rank of J is a direct summand of J if, and only if,*

(a) *J/C is separable;*

(b) *all the elements $\neq 0$ of J/C have the same genus in J/C.*

Proof. C is, as a complete subgroup of J, a direct summand of J, i.e., $J = C + J'$. Suppose now that every closed subgroup of finite rank of J is a direct summand of J. Either J' is of rank one and conditions (a) and (b) are obvious, or there exist two different rational closed subgroups S and T of J'. The closed subgroup U of J', generated by S and T, is a direct summand of J and of J', and S and T are direct summands of J and of U. By Corollary 2.9 either S and T or S and U/T are isomorphic. In the latter case $U = S' + T$, where S and S' are isomorphic. If $s \neq 0$ is an element of S', $t \neq 0$ an element of T, the closed subgroup of U, generated by $s + t$, is a direct summand of U and therefore $| s + t < U | = | S |$ or $= | T |$. Since the genus of $s + t$ is the g.c.d. of the genera of s and of t, this implies that either $| S | \leqq | T |$ or $| T | \leqq | S |$. Since S and T are not complete, and since the same argument applies to every

[8] The analogous theorem for primary abelian groups may be mentioned. The primary abelian group P does not contain elements of infinite height if, and only if, every finite subset of P is contained in a finite direct summand of P.

element $s + nt$ and $ns + t$ with positive integer n, it follows from condition 3.1 that S and T have the same genus, i.e., that J satisfies (a) and (b).

Suppose now that J satisfies the conditions (a) and (b) that S is a closed subgroup of finite rank in J, and that S' is its greatest complete subgroup. Then S' is a direct summand of C, i.e., $C = C' + S'$ and $S = S'' + S'$, where S'' is the intersection of S and $J' + C'$. The subgroup S^* of J/C which contains all those classes, containing elements of S, is exactly represented by S'' and by Corollary 4.4 is a direct summand of $J^* = J/C = S^* + T^*$. If T is the subgroup of J', representing T^*, then $J = T + S'' + S' + C'$, i.e., S is a direct summand of J.

LEMMA 4.6. *If J is separable and S a direct summand of finite rank of J, then J/S is separable.*

Proof. If F^* is a finite subset of $J^* = J/S$, F a subset of J, representing F^*, then J has by Corollary 4.3 a completely reducible direct summand D of finite rank which contains S and F. Since S is a direct summand of J, S is also a direct summand of D, i.e., $D = S + T$ and T is completely reducible by Corollary 3.5. The subgroup T^* of J^*, represented by T, is a completely reducible direct summand of J^* and contains F^*, i.e., J/S is separable.

THEOREM 4.7. *Every countable separable group is completely reducible.*

Proof. Let $b_1, b_2, \cdots, b_i, \cdots$ be an enumeration of the elements of the countable separable group J. It follows by complete induction from Corollary 4.3 that there exist completely reducible direct summands J_i of J such that $r(J_i)$ is finite, J_{i-1} is a direct summand of J_i, and the elements b_j with $j \leqq i$ are contained in J_i. Then $J_i = S_i + J_{i-1}$, S_i is completely reducible by Corollary 3.5 and J the direct sum of the completely reducible groups $J_1, S_2, S_3, \cdots$, i.e., J is completely reducible.

Chapter III. Types of elements and subgroups

5. Types of elements in separable groups.

DEFINITION 5.1. *The element b of the group J is a primitive element of genus* $\mathbf{s}$ *(in J), if*

$$b \equiv 0 \bmod (J, \mathbf{s} \leqq |x|), \qquad b \not\equiv 0 \bmod (J, \mathbf{s} < |x|),$$

$$m(b < J) = m(b < J/(J, \mathbf{s} < |x|)) = m(b < J(\mathbf{s})^*).$$

The subset F of J is primitive (in J), if F is finite, its elements are primitive elements in J and different elements of F have different genus in J.

LEMMA 5.2. *The finite subset F of the separable group J is a primitive set in J if, and only if,*

(a) *different elements of H have different genus in J;*

(b) *the closed subgroup $\bar{F}$ of J, generated by F, is a direct summand of J and F is a basis of $\bar{F}$.*

Proof. If the conditions (a) and (b) are satisfied, then every element in F is $\neq 0$ and generates a closed subgroup of J which is a direct summand of J, i.e., every element of F is primitive and F is therefore a primitive set.

If b is a primitive element in the separable group J, then the closed subgroup of J, generated by b, is by Theorem 4.2 a direct summand of J. If F is a primitive set in the separable group J, then F contains as a finite set an element b such that $|f| \not< |b|$ for every element f of F. Then $J = \bar{b} + J'$, where $\bar{b}$ is the closed subgroup of J, generated by b, and the elements $\neq b$ of F are all contained in J. Since by Lemma 4.6 also J' is separable, and since elements of J' are primitive in J' if, and only if, they are primitive in J, (b) follows by complete induction from the proved facts.

COROLLARY 5.3. *Suppose that the sets $b_1, \cdots, b_k$ and $b_1', \cdots, b_k'$ are primitive sets in the separable group J. Then there exists a (proper) automorphism α of J such that $b_i\alpha = b_i'$ for every i if, and only if, $m(b_i < J) = m(b_i' < J)$ for every i.*

Proof. Since J is separable, there exists a completely reducible direct summand D of finite rank which contains all the elements b_i and b_i'. By Lemma 5.2

$$D = \sum_{i=1}^{k} \bar{b}_i + B = \sum_{i=1}^{k} \bar{b}_i' + B',$$

where $\bar{b}$ is the closed subgroup of D, generated by b. Since b_i and b_i' have the same multiplicity and therefore the same genus, and since D is a completely reducible group of finite rank, Theorem 2.8 and Corollary 2.9 imply that B and B' are isomorphic. Since $\bar{b}_i$ and $\bar{b}_i'$ are isomorphic, and since $J = D + J'$, there exists an automorphism α of J such that $b_i\alpha = b_i'$ for every i, $u\alpha = u$ for u in J', $B\alpha = B'$, and this proves the corollary.

Another consequence of Lemma 5.2 is

COROLLARY 5.4. *If b is a primitive element of genus $\mathbf{s}$ in the separable group J, $|b| \neq |c|$ and $m(b < J) \mid m(c < J)$, then $b + c$ is also a primitive element of genus $\mathbf{s}$.*

DEFINITION 5.5. *$\mathbf{s}$ is a regular or singular genus in the group J according as $1 \leqq r(J(\mathbf{s})^*)$.*

LEMMA 5.6. *Suppose that b is a primitive element of genus $\mathbf{s}$ in the separable group J.*

(a) *If $\mathbf{s}$ is singular, the primitive elements of genus $\mathbf{s}$ are exactly the elements $r(b + b')$, where $b' \equiv 0 \bmod (J, \mathbf{s} < |x|)$, $m(b < J) \mid m(b' < J)$ and r is a rational number $\neq 0$ such that rb exists in J.*

(b) *If $\mathbf{s}$ is regular, there exists a (primitive) element b' of genus $\mathbf{s}$ in J such that b, b' is a basis of a direct summand of J.*

Proof. If b is a primitive element of genus $\mathbf{s}$ in the separable group J, then $J = \bar{b} + J'$, where $\bar{b}$ is the closed subgroup of J generated by b (Lemma 5.2). If b' is an element of $(J, \mathbf{s} < |x|)$, then b' is contained in J', and if $m(b < J) \mid m(b' < J)$, then $m(b + b' < J) = m(b < J)$ and $b \equiv b + b' \bmod(J, \mathbf{s} < |x|)$, i.e., $b + b'$ and every $r(b + b') \neq 0$ in J are primitive elements of genus $\mathbf{s}$ in J.

If $\mathbf{s}$ is regular, then J' contains an element b'' such that $b'' \equiv 0 \bmod(J', \mathbf{s} \leqq |x|)$, $b'' \not\equiv 0 \bmod(J', \mathbf{s} < |x|)$. Since J' is by Lemma 4.6 also separable, b'' is contained in a completely reducible direct summand D of finite rank for J' and D has a rational direct summand of genus $\mathbf{s}$, since $(D, \mathbf{s} \leqq |x|) \neq (D, \mathbf{s} < |x|)$, and thus (b) is proved.

If finally $\mathbf{s}$ is singular and b'' a primitive element of genus $\mathbf{s}$ in J, then $b'' = rb + c$, where $rb \neq 0$, is an element of $\bar{b}$ and c an element of J'. Since $\mathbf{s} = |b''| = |rb|$, it follows that $\mathbf{s} \leqq |c|$, and since $(J', \mathbf{s} \leqq |x|) = (J', \mathbf{s} < |x|) = (J, \mathbf{s} < |x|)$, c is an element of $(J, \mathbf{s} < |x|)$. Since $b'' \equiv rb \bmod (J, \mathbf{s} < |x|)$, i.e., $m(b'' < J) = m(rb < J)$ as b'' is a primitive element of genus $\mathbf{s}$, $m(rb < J) / m(c < J)$ and therefore $c' = r^{-1}c$ exists in J and $b'' = r(b + c')$ has the required form.

We shall employ the notation

$$m(b, \mathbf{s}) = m(b < J, \mathbf{s}) \quad = m(b < J/(J, |x| \not< \mathbf{s})),$$

$$m(b, \mathbf{s}+) = m(b < J, \mathbf{s}+) = m(b < J/(J, |x| \not\leqq \mathbf{s})).$$

LEMMA 5.7. *Suppose that $b_1, \cdots, b_k$ is a primitive set in the separable group J and that $b = \sum_{i=1}^{k} b_i$.*

(a) *If $\mathbf{s} \neq |b_i|$ for every i, then $m(b, \mathbf{s}) = m(b, \mathbf{s}+)$.*

(b) *If $|b_j| \not< |b_i| = \mathbf{s}$ for every j, then*

$$m(b < J, \mathbf{s}) = \infty, \qquad m(b < J, \mathbf{s}+) = m(b_i < J).$$

(c) *If $|b_j| < |b_i| = \mathbf{s}$ for some j, then*

$$m(b < J, \mathbf{s}) \quad = \text{g.c.d. of all } m(b_j < J) \text{ with } |b_j| < \mathbf{s};$$

$$m(b < J, \mathbf{s}+) = \text{g.c.d. of all } m(b_j < J) \text{ with } |b_j| \leqq \mathbf{s};$$

$$m(b < J, \mathbf{s}) \neq \infty, \qquad m(b < J, \mathbf{s}+) / m(b < J, \mathbf{s})$$

and $q(b, \mathbf{s}) = m(b < J, \mathbf{s}) \div m(b < J, \mathbf{s}+)$ is an ordinary positive integer; $m(b_i < J)$ and $m(b < J, \mathbf{s}+)$ have the same finite p-value for every prime divisor p of $q(b, \mathbf{s})$.

Proof. If $\mathbf{s}$ is any genus, then $b \equiv \sum_{\mathbf{s} < |b_i|} b_i \equiv \sum_{\mathbf{s} \leqq |b_i|} b_i \bmod (J, |x| \not< \mathbf{s})$. Since the elements b_i form a primitive set in the separable group J, it follows from Lemma 5.2 that $m(b, \mathbf{s})$ and $m(b, \mathbf{s}+)$ have the values given in the lemma. The other statements are consequences of these evaluations and of the fact that $|u| < |v|$ implies the existence of a positive integer h such that $m(u < J) / hm(v < J)$.

NOTATIONS. *If $b \neq 0$ is any element of the separable group J, $\mathbf{s}$ a genus such that $m(b < J, \mathbf{s}) \neq \infty$, $q(b, \mathbf{s})$ therefore an ordinary positive integer, and g any (generalized) number, then*

$(g, b, \mathbf{s})^*$ is the greatest divisor of g which is relatively prime to $q(b, \mathbf{s})$;

$$(g, b, \mathbf{s})^{**} = \text{g.c.d. of } (g, b, \mathbf{s})^* \text{ and } (m(b, \mathbf{s}), b, \mathbf{s})^*;^{9}$$

$$(g, b, \mathbf{s}) = (g, b, \mathbf{s})^* \div (g, b, \mathbf{s})^{**}.$$

[9] $(g, b, \mathbf{s})^{**} = [\text{g.c.d. of } (g, b, \mathbf{s})^* \text{ and } m(b, \mathbf{s})] = [\text{g.c.d. of } (g, b, \mathbf{s})^* \text{ and } m(b, \mathbf{s}+)]$ by Lemma 5.7.

Note that by Lemma 5.7

$$(m(b, \mathbf{s}), b, \mathbf{s})^* = (m(b, \mathbf{s}+), b, \mathbf{s})^*.$$

LEMMA 5.8. *Suppose that, for every genus* $\mathbf{s}$, $r(\mathbf{s})$ *is a number of genus* $\mathbf{s}$. *Then there exists corresponding to every element* $b \neq 0$ *of the separable group* J *a primitive set* $b_1, \cdots, b_k$ *in* J *such that*

(a) $$b = \sum_{i=1}^{k} b_i ;$$

(b) $m(b < J, \mathbf{s}) \neq m(b < J, \mathbf{s}+)$ *if, and only if, (exactly) one of the elements* b_i *is of genus* $\mathbf{s}$;

(c) $m(b < J, \mathbf{s}) = \infty$, $m(b < J, \mathbf{s}+) = m(b_i < J)$, *if* $|b_i| = \mathbf{s}$ *and* $m(b < J, \mathbf{t}) = m(b < J, \mathbf{t}+)(= \infty)$ *for* $\mathbf{t} < \mathbf{s}$.

(d) *If* $\mathbf{s}$ *is a regular genus,* $|b_i| = \mathbf{s}$ *and* $m(b < J, \mathbf{s}) \neq \infty$, *then*

$$m(b_i < J) = m(b, \mathbf{s}+)(r(\mathbf{s}), b, \mathbf{s}).^{10}$$

(e) *If* $\mathbf{s}$ *is a singular genus,* $|b_i| = \mathbf{s}$ *and* $m(b < J, \mathbf{s}) \neq \infty$, *then*

$$m(b_i < J) = m(b, \mathbf{s}+)(r(\mathbf{s}), b, \mathbf{s})h_i,^{10}$$

where h_i *is an ordinary positive integer, relatively prime to* $q(b, \mathbf{s})$ *and chosen at random in its class* $H(h_i, b, \mathbf{s}) = H(b_1, \cdots, b_k, b, \mathbf{s})$, *where* $H(h, b, \mathbf{s})$ *denotes the smallest class of ordinary integers which contains* h_i *complete classes of residues* mod $q(b, \mathbf{s})$ *and with* n *also every* nn' *for* $n' \mid r(\mathbf{s})_\infty$, n' *an ordinary integer.*

Proof. 1. Since J is a separable group, every element $b \neq 0$ of J is contained in a completely reducible direct summand D of J. There exists a smallest partial reduction of D and, if $b(\mathbf{s})$ is the component of b in the summand, containing the elements of genus $\mathbf{s}$, the elements $b(\mathbf{s}) \neq 0$ are primitive elements of genus $\mathbf{s}$ and form therefore a primitive set (in D and in J) whose sum is b, i.e., to every element $b \neq 0$ of J there exists a primitive set in J whose sum is b.

2. If the elements $w, b_1, \cdots, b_k$ form a primitive set in J whose sum is b, if $|w| = \mathbf{s}$ and $m(b, \mathbf{s}) = m(b, \mathbf{s}+)$, it follows from Lemma 5.7 that $m(b, \mathbf{s}) \neq \infty$; that there exist elements b_i, say $b_1, \cdots, b_h$, such that $|b_j| < \mathbf{s}$ if, and only if, $1 \leqq j \leqq h$, $0 < h \leqq k$; that

$$c_j = m(b_j) \div \{\text{g.c.d. } m(b_j), m(w)\}$$

is for $1 \leqq j \leqq h$ an ordinary positive integer; and that 1 is the g.c.d. of $c_1, \cdots, c_h$. There exist therefore integers c_j' such that $\sum_{j=1}^{h} c_j' c_j = 1$ and a primitive set in J is defined by

$$b_i' = b_i + c_i' c_i w \qquad (1 \leqq i \leqq h),$$

$$b_i' = b_i \qquad (h < i),$$

[10] $m(b, \mathbf{s}+)(r(\mathbf{s}), b, \mathbf{s})$ = l.c.m. of $m(b, \mathbf{s}+)$ and $(r(\mathbf{s}), b, \mathbf{s})^*$ = l.c.m. of $m(b, \mathbf{s}+)$ and of all divisors of $r(\mathbf{s})$ which are relatively prime to $q(b, \mathbf{s})$.

since $|b_i| < |w|$ and $m(b_i) / m(c_i w) = c_i m(w)$ for $1 \leqq i \leqq h$, i.e., b_i and b_i' are primitive elements of the same genus and the same multiplicity in J (Corollary 5.4). Since the sum of this new primitive set is b, and since this new primitive set contains fewer elements than the old, we have the following result.

A primitive set in J with sum b satisfies the condition (b) if, and only if, it contains as few elements as possible. There exist therefore to every element $b \neq 0$ in J primitive sets in J which satisfy (a) and (b).

3. Suppose now that the primitive set $b_1, \cdots, b_k$ satisfies the conditions (a) and (b), that $\mathbf{s}$ is the genus of b_i, and that $m(b < J, \mathbf{s}) \neq \infty$. Then some of the elements b_j satisfy $|b_j| < \mathbf{s}$ and it can be assumed that exactly the elements b_j with $1 \leqq j \leqq h$, $1 \leqq h$ satisfy this condition.

$(m(b, \mathbf{s}), b, \mathbf{s})^* = (m(b, \mathbf{s}+), b, \mathbf{s})^*$ is by Lemma 5.7 the g.c.d. of the numbers $(m(b_j), b, \mathbf{s})^*$ with $1 \leqq j \leqq h$ and therefore a true divisor of $(m(b_i), b, \mathbf{s})^*$. Then

$$d_i = (m(b_i), b, \mathbf{s})^* \,\dot{\div}\, (m(b, \mathbf{s}), b, \mathbf{s})^*$$

is relatively prime to $q(b, \mathbf{s})$.

The numbers

$$c_j = m(b_j) \,\dot{\div}\, \{\text{g.c.d. } m(b_j), m(b_i)\} \qquad (1 \leqq j \leqq h)$$

are ordinary positive integers whose g.c.d. is $q(b, \mathbf{s})$.

Suppose now that w is an ordinary positive integer, dividing the finite part $(d_i)_f$ of d_i. Then there exists a decomposition

$$w = \prod_{j=1}^{h} w_j$$

such that w_j and $w_{j'}$ are relatively prime for $j \neq j'$ and such that $w_j / m(b_i) \,\dot{\div}\, \{\text{g.c.d. } m(b_j), m(b_i)\}$, since w is relatively prime to $q(b, \mathbf{s})$ and since $m(b, \mathbf{s}+)$ is the g.c.d. of the numbers g.c.d. $(m(b_j), m(b_i))$ with $1 \leqq j \leqq h$.

Then c_j and w_j are relatively prime integers and the g.c.d. of the numbers $c_j w w_j^{-1}$ with $1 \leqq j \leqq h$ is exactly $q(b, \mathbf{s})$. There exist therefore integers c_j' such that

$$q(b, \mathbf{s}) = \sum_{j=1}^{h} c_j' c_j (w w_j^{-1}) \text{ or } q(b, \mathbf{s}) w^{-1} = \sum_{j=1}^{n} c_j' c_j w_j^{-1} \qquad (\not\equiv 0 \bmod 1).$$

Since $|b_j| < \mathbf{s}$ and $w_j m(b_j) / c_j m(b_i)$ for $1 \leqq j \leqq h$, a primitive set is defined by

$$b_j' = b_j \qquad (h < j \neq i),$$

$$b_j' = b_j + n c_j' c_j w_j^{-1} b_i \qquad (1 \leqq j \leqq h),$$

$$b_j' = (1 - n q(b, \mathbf{s}) w^{-1}) b_i \qquad (j = i),$$

where n is any given integer. This primitive set satisfies (a), (b), and

$$m(b_j') = m(b_j) \qquad (j \neq i),$$

$$m(b_j') = |w - n q(b, \mathbf{s})| \, w^{-1} m(b_i) \qquad (i = j).$$

Since it is possible to apply this construction[11] upon several indices i without changing the effect on the other indices, we have the following result.

(5.8.3) Suppose that the elements b_i form a primitive set, satisfying (a) and (b) and that for $m(b, | b_i |) \neq \infty$ n_i is any integer and $w_i / d_i = (m(b_i), b, | b_i |)^* \vdots (m(b, | b_i |), b, | b_i |)^*$. Then there exists a primitive set b_i' which satisfies (a) and (b) and

$$m(b_i') = m(b_i), \text{ if } m(b, | b_i |) = \infty,$$

$$= | w_i - n_i q(b, | b_i |) | w_i^{-1} m(b_i), \text{ if } m(b, | b_i |) \neq \infty.$$

4. Suppose that $b_1, \cdots, b_k$ is a primitive set, satisfying (a) and (b), and that the indices i, h and the numbers w, w_j, c_j, c_j' for $1 \leqq j \leqq h$ and d_i have the same meaning as during the proof of (5.8.3). Suppose furthermore that the genus $\mathbf{s}$ of b_i is regular for the group J. Then there exists by Lemma 5.6b an element e in J such that b_i, e is a basis of a direct summand of J and satisfies $m(b_i < J) = m(e < J)$. Since $| b_j | < \mathbf{s}$ and $w_j m(b_j) / c_j m(b_i) = c_j m(e)$, and since e is a primitive element of genus $\mathbf{s}$ in J, it follows from Lemma 5.2 and Corollary 5.3 that a primitive set in J is defined by

$$b_j' = b_j \qquad (h < j \neq i),$$

$$b_j' = b_j + c_j c_j' w_j^{-1} e \qquad (1 \leqq j \leqq h),$$

$$b_j' = b_i - w^{-1} q(b, \mathbf{s}) e \qquad (j = i).$$

This primitive set satisfies (a), (b) and

$$m(b_j') = m(b_j) \qquad (j \neq i),$$

$$m(b_j') = [\text{g.c.d. of } m(b_i) \text{ and } w^{-1} q(b, \mathbf{s}) m(e)] = w^{-1} m(b_i) \qquad (j = i),$$

since $m(b_i) = m(e)$ and since w and $q(b, \mathbf{s})$ are relatively prime.

Since it is possible to apply this construction upon several indices i without changing the effect on the other indices, we have the following result.

(5.8.4) Suppose that the elements $b_1, \cdots, b_k$ form a primitive set which satisfies (a) and (b), and for every i such that $m(b, | b_i |) \neq \infty$ and $| b_i |$ is a regular genus for J the integer $w_i/d_i = (m(b_i), b, | b_i |)^* \vdots (m(b, | b_i |), b, | b_i |)^*$. Then there exists a primitive set $b_1', \cdots, b_k'$ which satisfies (a) and (b) and $m(b_i') = m(b_i)$, if $m(b, | b_i |) = \infty$ or $| b_i |$ is a singular genus for J, or $m(b_i') = w_i^{-1} m(b_i)$, if $m(b, | b_i |) \neq \infty$ and $| b_i |$ is a regular genus.

5. Suppose that $b_1, \cdots, b_k$ is a primitive set which satisfies (a) and (b) and that $m(b, | b_i |) \neq \infty$ if, and only if, $1 \leqq i \leqq z$, $0 \leqq z \leqq k$. Since $r_i = m(b, | b_i | +)(r(| b_i |), b, | b_i |)$ has the genus $| b_i |$ for $1 \leqq i \leqq z$, there exists an ordinary positive integer n_i' which is relatively prime to $q(b, | b_i |)$ such that $r_i / n_i' m(b_i)$ and there exist integers n_i, n_i'' such that $1 = n_i q(b, | b_i |) + n_i' n_i''$. There exists therefore by (5.8.3) a primitive set $b_1', \cdots, b_k'$ which satisfies (a)

[11] Since $m(rb_i) = m(b_i)$, if numerator and denominator of r are divisors of the infinite part of $m(b_i)$, it suffices for this construction to assume that w is a divisor of d_i.

and (b) and

$$m(b_i') = m(b_i) \qquad (z < i),$$

$$m(b_i') = | n_i' n_i'' | m(b_i) \qquad (1 \leqq i \leqq z).$$

Thus it follows from Lemma 5.7 that

(5.8.5) There exists a primitive set $b_1, \cdots, b_k$ which satisfies (a), (b), (c) and the additional condition (d*).

(d*) If $m(b, | b_i |) \neq \infty$, then

$$m(b_i) = m(b, | b_i | +)(r(| b_i |), b, | b_i |)h_i',$$

where h_i' is an ordinary positive integer which is relatively prime to $q(b, | b_i |)$.

6. Since the numbers h_i' in (5.8.5d*) satisfy the condition imposed on the numbers w_i in (5.8.3), it follows that these numbers h_i' can be chosen at random in their class of residues mod $q(b, | b_i |)$; and since $m(b_i) = pm(b_i)$, if p is a prime number which divides the infinite part $r(| b_i |)_\infty$ of $m(b_i)$, it follows that h_i' can be chosen at random in its class $H(h_i', b, | b_i |)$.

7. If $m(b, | b_i |) \neq \infty$ and $| b_i |$ is a regular genus, then the numbers h_i' in (5.8.5d*) satisfy the condition imposed upon the numbers w_i in (5.8.5), and the existence of a primitive set which satisfies (a) to (e) follows therefore from (5.8.5) and the facts proved in heading 6.

Let $\mathbf{s}$ be a singular genus for the separable group J; n an ordinary positive integer which is relatively prime to the infinite part of the numbers of genus $\mathbf{s}$; $Y = Y(J, n, \mathbf{s})$ the subgroup of J generated by nJ and $(J, | x | \leqq\!\!\!\!/\; \mathbf{s})$; and $Z = Z(J, n, \mathbf{s})$ the group of those classes of $J/Y(J, n, \mathbf{s})$ which contain elements of $(J, | x | \nless \mathbf{s})$.

Since J is a separable group and $\mathbf{s}$ a singular genus, it follows from Theorem 2.8 that there exists a rational subgroup R of genus $\mathbf{s}$ and a subgroup J' such that

$$J = R + J',\ (J, | x | \nless \mathbf{s}) = R + (J, | x | \leqq\!\!\!\!/\; \mathbf{s}),\ (J, \mathbf{s} \leqq | x |) = R + (J, \mathbf{s} < | x |),\ (J', | x | \nless \mathbf{s}) = (J', | x | \leqq\!\!\!\!/\; \mathbf{s}),\ (J', \mathbf{s} < | x |) = (J', \mathbf{s} \leqq | x |).$$

Since every class of $Z(J, n, \mathbf{s})$ contains elements of R, i.e., primitive elements of genus $\mathbf{s}$ and since every primitive element of genus $\mathbf{s}$ is contained in a class of Z, it follows that $Z(J, n, \mathbf{s})$ and R/nR are isomorphic, i.e., since n is relatively prime to the infinite part of the numbers of genus $\mathbf{s}$, $Z(J, n, \mathbf{s})$ *is a cyclic group of the finite order* n.

nJ, $(J, | x | \leqq\!\!\!\!/\; \mathbf{s})$ and $(J, | x | \nless \mathbf{s})$ are characteristic subgroups of J, and every (proper) automorphism of J induces therefore an automorphism of $Z(J, n, \mathbf{s})$. Two elements z and z' of Z are called *conjugate*, if there exists an automorphism of J which maps z upon z'. The elements of Z are thus distributed into classes of conjugate elements.

LEMMA 5.9. *Suppose that* $\mathbf{s}$ *is a singular genus of the separable group* J *and that the positive integer* n *is relatively prime to the infinite part of the numbers of genus* $\mathbf{s}$.

(a) *The elements z and z' of $Z(J, n, \mathbf{s})$ are conjugate if, and only if, there exist positive integers h,h' which are divisors of the infinite part of the numbers of genus* $\mathbf{s}$ *and which satisfy $hz = h'z'$.*

(b) *The elements z and z' of $Z(J, n, \mathbf{s})$ are conjugate if, and only if, the primitive elements of genus* $\mathbf{s}$, *representing z, have the same multiplicities as those representing z'.*

(c) *The primitive elements b and b' of genus* $\mathbf{s}$ *represent conjugate elements of $Z(J, n, \mathbf{s})$ if, and only if,*

(c′) *b and b' have the same order q mod $Y(J, n, \mathbf{s})$;*

(c″) *there exist ordinary integers k and k' which are relatively prime to q and satisfy*

$$k \equiv k' \bmod q, \qquad m(kb < J) = m(k'b' < J).$$

Proof. If h, h' are ordinary positive integers, dividing the infinite part of the numbers of genus $\mathbf{s}$, and if z and z' are elements of $Z(J, n, \mathbf{s})$ such that $hz = h'z'$, then z' contains a primitive element b' of genus $\mathbf{s}$ and the primitive element $b = hh'^{-1}b'$ of genus $\mathbf{s}$ in z satisfies $m(b) = m(b')$, i.e., z and z' are represented by primitive elements of the same multiplicity.

If z and z' are elements of Z which are represented by primitive elements b and b', respectively, such that $m(b) = m(b')$, there exists by Corollary 5.3 an automorphism of J which maps b upon b' and therefore z upon z', i.e., z and z' are conjugate.

Suppose that z and z' are conjugate and that the primitive elements b and b' of genus $\mathbf{s}$ represent z and z' respectively. Then there exists an automorphism of J which maps z' upon z and therefore b' upon the element b'' representing z. Moreover, $m(b') = m(b'')$. Since b, b'' are primitive elements of the singular genus $\mathbf{s}$, it follows from Lemma 5.6a that

$$b'' = r(b + c)$$

where $m(b'') = m(rb)$, rb exists in J and $c \equiv 0 \bmod (J, \mathbf{s} < |x|)$. Since $b'' \equiv b$ mod $Y(J, n, \mathbf{s})$ and $rb \equiv r(b + c) \bmod Y$, this implies $b \equiv rb \bmod Y$ and therefore $b \equiv rb \bmod nJ$, $b \equiv rb \bmod n\bar{b}$, where $\bar{b}$ is the closed subgroup of J, generated by b. The order q of b mod Y is nn'^{-1} where n' is the g.c.d. of n and $m(b)$. Since b and rb have the same order mod Y, n' is also the g.c.d. of n and $m(rb)$ and therefore $n'^{-1}(r - 1)b \equiv 0 \bmod q\bar{b}$, q and $m(n'^{-1}b)$ are relatively prime. If $r = r'r''^{-1}$, where r' and r'' are relatively prime integers, then r'' is relatively prime to q (as a divisor of $m(n'^{-1}b)$) and $r' \equiv r'' \bmod q$ and furthermore $m(r'b) = m(r''b')$, i.e., b and b' satisfy the conditions (c′) and (c″).

Suppose now that the primitive elements b and b' of genus $\mathbf{s}$ satisfy the conditions (c′) and (c″). Since k is relatively prime to q, there exists an integer k'' such that $kk'' \equiv 1 \bmod q$ and since $b \equiv b' \equiv 0 \bmod nq^{-1}J$, it follows that $b \equiv kk''b \equiv k'k''b \bmod Y(J, n, \mathbf{s})$, $b' \equiv kk''b' \equiv k'k''b' \bmod Y$, and since $m(kb) = m(k'b')$, this implies that kb and $k'b'$ and therefore b and b' represent conjugate elements of Z.

If z and z' are conjugate elements of Z, b a primitive element of genus $\mathbf{s}$,

representing z, then there exist relatively prime integers h, h' such that $m(b) = m(hh'^{-1}b)$ and $hh'^{-1}b$ represents z'. Then h and h' divide the infinite part of the numbers of genus $\mathbf{s}$, and $hz = h'z'$. This completes the proof of the lemma.

NOTATION. If $b \neq 0$ is an element of the separable group J, $\mathbf{s}$ a singular genus of J such that $\infty \neq m(b, \mathbf{s}) \neq m(b, \mathbf{s}+)$, then $n(b, \mathbf{s}) = m(b, \mathbf{s}) \dot{:}_{\mathbf{s}} (m(b, \mathbf{s}), b, \mathbf{s})^*$.

COROLLARY 5.10. (a) *If $b \neq 0$ is an element of the separable group J, $\mathbf{s}$ a singular genus such that $\infty \neq m(b, \mathbf{s}) \neq m(b, \mathbf{s}+)$, then b represents a certain element of order $q(b, \mathbf{s})$ of $Z(J, n(b, \mathbf{s}), \mathbf{s})$ and therefore a class $C(b, \mathbf{s})$ of conjugate elements of $Z(J, n(b, \mathbf{s}), \mathbf{s})$.*

(b) *Suppose that, for every genus $\mathbf{t}$, $r(\mathbf{t})$ is a number of genus $\mathbf{t}$ and that the elements b and b' of the separable group J satisfy $\infty \neq m(b, \mathbf{s}) = m(b', \mathbf{s}) \neq m(b, \mathbf{s}+) = m(b', \mathbf{s}+)$ for a certain singular genus $\mathbf{s}$. Then $n(b, \mathbf{s}) = n(b', \mathbf{s})$ and $q(b, \mathbf{s}) = q(b', \mathbf{s})$. Moreover, $C(b, \mathbf{s}) = C(b', \mathbf{s})$ if, and only if, $H(b_1, \cdots, b_k, b, \mathbf{s}) = H(b_1', \cdots, b_{k'}', b', \mathbf{s})$, where $b_1, \cdots, b_k$ and $b_1', \cdots, b_{k'}'$, are primitive sets in J with sum b and b' respectively which satisfy the conditions* (b) to (e) *of Lemma 5.8. $H(b_1, \cdots, b_k, b, \mathbf{s}) = H(b, \mathbf{s})$ depends therefore only on b, $\mathbf{s}$, and the numbers $r(\mathbf{t})$, representing the genera $\mathbf{t}$, but not on the particular representation of b by a "canonical" primitive set.*

Proof. If $b \neq 0$ is an element of the separable group J, then there exists a primitive set $b_1, \cdots, b_k$, satisfying the conditions (a) to (e) of Lemma 5.8. If $\mathbf{s}$ is a singular genus such that $\infty \neq m(b, \mathbf{s}) \neq m(b, \mathbf{s}+)$, then exactly one of the elements b_i, say b_1 has the genus $\mathbf{s}$ in J. Then $b \equiv b_1 \bmod Y(J, n(b, \mathbf{s}), \mathbf{s})$ and now the corollary is a consequence of Lemma 5.8 and Lemma 5.9.

DEFINITION 5.11. *The subsets S and T of the group J are isotype (have the same type in J) if there exists a (proper) automorphism of J which maps S upon T.*

THEOREM 5.12. *The elements b and b' of the separable group J are isotype in J if, and only if,*[12]

(a) *$m(b, \mathbf{s}) = m(b', \mathbf{s})$ and $m(b, \mathbf{s}+) = m(b', \mathbf{s}+)$ for every genus $\mathbf{s}$;*

(b) *$C(b, \mathbf{s}) = C(b', \mathbf{s})$ for every singular genus $\mathbf{s}$ such that $\infty \neq m(b, \mathbf{s}) = m(b', \mathbf{s}) \neq m(b, \mathbf{s}+) = m(b', \mathbf{s}+)$ and therefore $n(b, \mathbf{s}) = n(b', \mathbf{s})$, $q(b, \mathbf{s}) = q(b', \mathbf{s})$.*

Proof. $m(b, \mathbf{s})$, $m(b, \mathbf{s}+)$ and $C(b, \mathbf{s})$ are invariants of the element b in the group J, since for their definitions only characteristic subgroups of J have been used. If the conditions (a) and (b) are satisfied by the elements b and b', there exists by Lemma 5.8 and by Corollary 5.10 a primitive set $b_1, \cdots, b_k$ with sum b and a primitive set $b_1', \cdots, b_k'$ with sum b' in J such that $m(b_i < J) = m(b_i' < J)$ for $1 \leqq i \leqq k$. There exists therefore by Corollary 5.3 an automorphism α of J such that $b_i\alpha = b_i'$, i.e., such that $b\alpha = b'$.

COROLLARY 5.13. *Suppose that b_i is an element of the separable group J_i and that J_1 and J_2 are isomorphic. Then there exists an isomorphism of J_1 upon J_2 which maps b_1 upon b_2 if, and only if,*

(a) *$m(b_1 < J_1, \mathbf{s}) = m(b_2 < J_2, \mathbf{s})$ and $m(b_1 < J_1, \mathbf{s}+) = m(b_2 < J_2, \mathbf{s}+)$ for every genus $\mathbf{s}$;*

[12] A reason for the so different nature of the invariants $m(b, \mathbf{s})$, $m(b, \mathbf{s}+)$ on the one hand and the invariants $C(b, \mathbf{s})$ on the other will be given in Corollary 6.11.

(b) *the primitive elements of genus* $\mathbf{s}$, *representing* $C(b_1, \mathbf{s})$, *have the same multiplicities as those representing* $C(b_2, \mathbf{s})$ *for every singular genus* $\mathbf{s}$ *of* J_1 *and of* J_2 *such that*

$$\infty \neq m(b_i < J_i, \mathbf{s}) \neq m(b_i < J_i, \mathbf{s}+).$$

For an isomorphism of J_2 upon J_1 maps b_2 upon an element b_1' such that b_1 and b_1' satisfy in J_1 the conditions (a) and (b) of Theorem 5.12 if, and only if, the conditions (a) and (b) of the Corollary 5.13 are satisfied. This follows from Lemma 5.9 and Corollary 5.10.

Corollary 5.14. *The elements* b *and* b' *of the separable group* J *satisfy the condition* (a) *of Theorem* 5.12 *if, and only if,*

(1) $m(b, \mathbf{s}) \neq m(b, \mathbf{s}+)$ *if, and only if,* $m(b', \mathbf{s}) \neq m(b', \mathbf{s}+)$;

(2) $\infty \neq m(b, \mathbf{s}) \neq m(b, \mathbf{s}+)$ *if, and only if,* $\infty \neq m(b', \mathbf{s}) \neq m(b', \mathbf{s}+)$;

(3) $m(b, \mathbf{s}+) = m(b', \mathbf{s}+)$, *if* $m(b, \mathbf{s}) \neq m(b, \mathbf{s}+)$.

Or (3) *can be replaced by*

(3′) $m(b, \mathbf{s}+) = m(b', \mathbf{s}+)$, *if* $\infty = m(b, \mathbf{s}) \neq m(b, \mathbf{s}+)$;

(3″) $q(b, \mathbf{s}) = q(b', \mathbf{s})$, *if* $\infty \neq m(b, \mathbf{s}) \neq m(b, \mathbf{s}+)$.

This follows from Lemma 5.7 and Lemma 5.8, since there exists but a finite number of genera $\mathbf{s}$ such that $m(b, \mathbf{s}) \neq m(b, \mathbf{s}+)$ and since $m(b, \mathbf{s}+) \neq m(b, \mathbf{s}) \neq \infty$ implies that $m(b, \mathbf{s})$ is the g.c.d. of those $m(b, \mathbf{t}+)$ with $\mathbf{t} < \mathbf{s}$ and $m(b, \mathbf{t}) \neq m(b, \mathbf{t}+)$.

6. Characteristic subgroups of separable groups.

Definition 6.1. *The subgroup* S *of the group* J *is a characteristic subgroup of* J, *if* S *is mapped upon itself by every proper automorphism of* J. S *is regular, if every proper or improper automorphism of* J *maps* S *upon a subgroup of* S. *A characteristic subgroup of* J *which is not regular is singular.*

Regular subgroups are usually called strictly characteristic, singular ones characteristic, but not strictly characteristic. The above notation has been adopted for brevity.

nJ and $(J, \mathbf{s} \leqq |x|)$ are examples of regular subgroups. Intersection and join of characteristic or regular subgroups are characteristic or regular subgroups.

Lemma 6.2. *Suppose that* $b \neq 0$ *is an element of the characteristic subgroup* S *of the separable group* J *and that the genus* $\mathbf{s}$ *satisfies* $m(b, \mathbf{s}) \neq m(b, \mathbf{s}+)$. *Then the element* w *of* J *is contained in* S, *if one of the following conditions is satisfied:*

1. $\mathbf{s} < |w|$, $m(b, \mathbf{s}+) / m(w)$.
2. $\mathbf{s} = |w|$ *and* $2m(b, \mathbf{s}+) / m(w)$
   or $m(b, \mathbf{s}+) / m(w)$ *and* S *is regular*
   or $\mathbf{s}$ *is regular*
   or $m(b, \mathbf{s}) \neq \infty$ *and* $q(b, \mathbf{s})$ *is odd.*

Two particular instances of this lemma will be needed for its proof.

(6.2.1) *Suppose that the characteristic subgroup* S *of the separable group* J *contains the sum* b *of the primitive set* $b_1, \cdots, b_k$ *in* J.

(a) $2b_i$ *is contained in* S.

(b) b_i *is contained in* S, *if* S *is regular or* $|b_i|$ *is regular.*

Proof. By Lemma 5.2 there exists a subgroup J' such that $J = \sum_{i=1}^{k} \bar{b}_i + J'$, where $\bar{b}_i$ is the closed subgroup of J, generated by b_i. A proper automorphism α of J is therefore defined by $b_j\alpha = b_j$ for $j \neq i$, $u\alpha = u$ for u in J', $b_j\alpha = -b_i$ for $j = i$, and an improper automorphism β is defined by $b_j\beta = b_j$ for $j \neq i$, $u\beta = u$ for u in J', $b_j\beta = 0$ for $j = i$. Then $b - b\alpha = 2b_i$ is contained in S. If S is regular, then $b - b\beta = b_i$ is contained in S. If finally $|b_i|$ is regular, there exists by Lemma 5.6b an element w in J' such that $m(w < J) = m(b_i < J)$ and $J' = \bar{w} + J''$, where $\bar{w}$ is the closed subgroup of J generated by w. Then a proper automorphism γ of J is defined by $b_j\gamma = b_j$ for $j \neq i$, $u\gamma = u$ for u in J', $b_j\gamma = b_i + w$ for $j = i$, and therefore $b - b\gamma = -w$ is contained in S. But since $-w$ and b_i are primitive elements of the same genus and the same multiplicity in J, it follows from Corollary 5.3 that b_i is contained in S.

(6.2.2) *If the characteristic subgroup S of the separable group J contains the primitive element b, then S contains every element c such that $m(b) / m(c)$.*

Proof. c is by Lemma 5.8 the sum of a primitive set $c_1, \cdots, c_k$ and Lemma 5.7 implies $m(b) / m(c) / m(c_i)$. If $|c_i| = |b|$, then there exists an ordinary positive integer h such that $hm(b) = m(hb) = m(c_i)$. Since hb is an element of S, it follows from Corollary 5.3 that c_i is contained in S. If $|b| < |c_i|$, then it follows from Corollary 5.4 that $b + c_i$ is a primitive element of the same genus and the same multiplicity as b and it follows therefore from Corollary 5.3 that $b + c_i$ and consequently c_i are elements of S.

(6.2.3) *If the assumptions:*

$b_1, \cdots, b_k$ *is a primitive set in the separable group J, complying with the conditions of Lemma* 5.8;

$b = \sum_{i=1}^{k} b_i$ *is an element of the characteristic subgroup S of J; the genus* $\mathbf{s}$ *of* b_1 *is singular for J;*

$\infty \neq m(b, \mathbf{s}) \neq m(b, \mathbf{s}+)$;

$m(b_1 < J) = m(b, \mathbf{s}+)fh$, *where h is an ordinary positive integer which is relatively prime to* $q(b, \mathbf{s})$;

are satisfied, then S contains $q(b, \mathbf{s})h^{-1}b_1$ and there exists an integer h' which is relatively prime to $q(b, \mathbf{s})$ such that $h^{-1}b_1 + h' \sum_{i=2}^{k} b_i$ is contained in S.

Proof. By Lemma 5.8 there exists a primitive set $b_1', \cdots, b_k'$ with sum b in J such that

$$m(b_i') = m(b_i) \qquad (i \neq 1),$$

$$m(b_i') = m(b, \mathbf{s}+)f(h + q(b, \mathbf{s})) \qquad (i = 1).$$

Corollary 5.3 implies therefore that $(h + q(b, \mathbf{s}))h^{-1}b_1 + \sum_{i=2}^{k} b_i$ and consequently $q(b, \mathbf{s})h^{-1}b_1$ are elements of S. Since h and $q(b, \mathbf{s})$ are relatively prime, there

exist integers h', q' such that $hh' + q'q(b, \mathbf{s}) = 1$ and

$$h^{-1}b_1 + h' \sum_{i=2}^{k} b_i = h'b + q'q(b, \mathbf{s})h^{-1}b_1$$

is an element of S.

Proof of Lemma 6.2. Suppose that $\mathbf{t}$ is the genus of w. Then it is possible to choose the number $r(\mathbf{u})$ of genus $\mathbf{u}$ in such a way that $r(\mathbf{u}) \mid m(w < J)$ for $\mathbf{u} \leqq \mathbf{t}$. There exists by Lemma 5.8 and (6.2.3) a primitive set $b_1, \cdots, b_k$ in J whose sum b' is an element of S such that $\mathbf{s}$ is the genus of b_1 and $m(b_1 < J) = m(b, \mathbf{s}+)$, if $m(b, \mathbf{s}) = \infty$, $m(b_1 < J) = m(b, \mathbf{s}+)(r(\mathbf{s}), b, \mathbf{s})$, if $m(b, \mathbf{s}) \neq \infty$.

If $\mathbf{s} < \mathbf{t}$ and $m(b, \mathbf{s}+) \mid m(w)$, then $m(b_1) \mid m(w)$ and $b_1 + w$ is therefore by Corollary 5.4 a primitive element of genus $\mathbf{s}$, satisfying $m(b_1) = m(b_1 + w)$. Corollary 5.3 implies therefore that $b' + w$ and consequently w are elements of S.

If $\mathbf{s}$ is regular or S is regular, then b_1 is by (6.2.1) an element of S. If $m(b, \mathbf{s}+) \mid m(w)$, then $m(b_1) \mid m(w)$ and w is therefore by (6.2.2) an element of S.

$2b_1$ is by (6.2.1) an element of S. If $2m(b, \mathbf{s}+) \mid m(w)$, then $m(2b_1) \mid m(w)$ and w is therefore by (6.2.2) an element of S.

COROLLARY 6.3. *If the primitive set* $b_1, \cdots, b_k$ *in the separable group* J *is a subset of the characteristic subgroup* S *of* J *and if the element* $b \neq 0$ *in* J *satisfies* l.c.m. $(|b_1|, \cdots, |b_k|) \leqq |b|$, g.c.d. $(m(b_1), \cdots, m(b_k)) \mid m(b)$, *then* b *is an element of* S.

Proof. From the assumptions there exist relatively prime positive integers h_i such that $m(b_i) \mid h_i m(b) = m(h_i b)$. Since b_i is an element of S, (6.2.2) implies that $h_i b$ is an element of S. Since the h_i are relatively prime, there exist integers h_i' such that $\sum_{i=1}^{k} h_i' h_i = 1$ and $b = \sum_{i=1}^{k} h_i' h_i b$ is an element of S.

DEFINITION 6.4. *If* S *is a subgroup of the separable group* J, $\mathbf{s}$ *a genus, then* $g(S, \mathbf{s}) = g(S < J, \mathbf{s})$ *is the* g.c.d. *of all the multiplicities* $m(b < J)$ *of primitive elements* b *of genus* $\mathbf{s}$ *in* J *which are elements of* S.

Note that $g(S, \mathbf{s}) = \infty$, if S does not contain primitive elements of genus $\mathbf{s}$.

THEOREM 6.5. (a) *If* S *is a characteristic subgroup of the separable group* J, *then* S *contains every primitive element* b *of genus* $\mathbf{s}$ *in* J *such that* $g(S, \mathbf{s}) \mid m(b < J)$.

(b) *If* S *is a regular subgroup of the separable group* J *and* T *a characteristic subgroup of* J, *then* $S \leqq T$ *if (and only if)* $g(T, \mathbf{s}) \mid g(S, \mathbf{s})$ *for every genus* $\mathbf{s}$.

(c) *If* S *and* T *are regular subgroups of the separable group* J, *then* $S = T$ *if (and only if)* $g(T, \mathbf{s}) = g(S, \mathbf{s})$ *for every genus* $\mathbf{s}$.

Proof. $m(b < J)$ has the same infinite part for every primitive element b of genus $\mathbf{s}$ in J. If the prime number p does not divide the infinite part of the numbers of genus $\mathbf{s}$, and if $g(S, \mathbf{s}) \neq \infty$, there exists a primitive element b_p of genus $\mathbf{s}$ in J which is an element of S such that the p-values of $m(b_p < J)$ and $g(S, \mathbf{s})$ are equal (and finite). Let now b be a primitive element of genus $\mathbf{s}$ in J such that $g(S, \mathbf{s}) \mid m(b < J)$. S contains a primitive element b' of genus $\mathbf{s}$. If $m(b' < J) \mid m(b < J)$, then $m(b < J) = m(hb' < J)$ for some integer $h \neq 0$,

and b is an element of S by Corollary 5.3. If $m(b' < J)$ is not a divisor of $m(b < J)$, there exists a positive integer k which is relatively prime to the infinite part of $m(b)$ such that $m(b' < J) / m(kb < J) = km(b < J)$. kb is therefore an element of S. Let h be the smallest positive integer such that hb is an element of S. Then h / k. Furthermore there exists in S by Corollary 5.3 a multiple of b whose multiplicity has the same p-value as $m(b_p < J)$. If $k \neq 1$ and p / k, there exists in S an element $wp^{-i}b$, where w is a positive integer relatively prime to p. If $w'w + k'k =$ g.c.d. of w and k, then $(w'w + k'k)b$ is an element of S and $0 < w'w + k'k < k$. That being impossible, $k = 1$, and b is an element of S.

If S is regular and the sum b of the primitive set $b_1, \cdots, b_k$ an element of S, then by (6.2.1) every b_i is contained in S. If the characteristic subgroup T of J satisfies $g(T, \mathbf{s}) / g(S, \mathbf{s})$, it follows from (a) that $S \leqq T$, i.e., (b) is true. (c) is a consequence of (b).

THEOREM 6.6. *Let, for every genus* $\mathbf{s}$, $g(\mathbf{s})$ *be either a (generalized) number or the symbol* ∞. *Then there exists a regular subgroup* S *of the separable group* J *such that* $g(S < J, \mathbf{s}) = g(\mathbf{s})$ *for every genus* $\mathbf{s}$ *if, and only if,*

(a) $g(\mathbf{s}) = \infty$, *if* $(J, \mathbf{s} < |x|) = (J, \mathbf{s} \leqq |x|)$;

(b) *if* $g(\mathbf{s}) \neq \infty$, *then*

(b1) $|g(\mathbf{s})| \leqq \mathbf{s}$;

(b2) *the infinite part of* $g(\mathbf{s})$ *is the infinite part of the numbers of genus* $\mathbf{s}$;

(b3) *the finite part* $g(\mathbf{s})_f / g(\mathbf{t})$ *for every* $\mathbf{t} < \mathbf{s}$;

(c) *if* $g(\mathbf{t}) \neq \infty$, $\mathbf{t} < \mathbf{s}$ *and* $(J, \mathbf{s} < |x|) \neq (J, \mathbf{s} \leqq |x|)$, *then* $g(\mathbf{s}) \neq \infty$.

Proof. A. Suppose that S is a regular subgroup of the separable group J. If $(J, \mathbf{s} < |x|) = (J, \mathbf{s} \leqq |x|)$, there do not exist primitive elements of genus $\mathbf{s}$ in J, i.e., $g(S, \mathbf{s})$ satisfies (a). That $g(S, \mathbf{s})$ satisfies (b1) and (b2) is obvious. (b3) and (c) are consequences of Corollary 6.3.

B. Suppose that the numbers $g(\mathbf{s})$ satisfy the conditions (a) to (c). Then put

$$S(\mathbf{s}) = (J; \mathbf{s} \leqq |x|, g(\mathbf{s})/m(x < J)).$$

Thus for every genus $\mathbf{s}$ a regular subgroup of J has been defined. $S(\mathbf{s}) = 0$ if, and only if, $g(\mathbf{s}) = \infty$ (by (c)). $S(\mathbf{s})$ consists exactly of the elements $x \neq 0$ with $\mathbf{s} \leqq |x|$, $g(\mathbf{s}) / m(x)$. (a), (b1), (b2) and (c) imply that $g(S(\mathbf{s}) < J, \mathbf{s}) = g(\mathbf{s})$.

Now let S be the subgroup of J, generated by all these subgroups $S(\mathbf{s})$. As the join of regular subgroups, S is a regular subgroup of J. Since by Lemma 5.8 every element of $S(\mathbf{s})$ is the sum of primitive elements b such that $\mathbf{s} \leqq |b|$, $g(\mathbf{s}) / m(b)$, it follows that every primitive element s of genus $\mathbf{t}$ which is contained in S has the form $s = u + v + w$, where

$$u = \sum_i u_i,\ u_i \text{ primitive, } |u_i| < \mathbf{t},\ g(|u_i|) / m(u_i);$$

$$v = \sum_i v_i,\ v_i \text{ primitive of genus } \mathbf{t},\ g(\mathbf{t}) / m(v_i);$$

$$w = \sum_i w_i,\ w_i \text{ primitive, } |w_i| \nleqq \mathbf{t},\ g(|w_i|) / m(w_i).$$

Then $s \equiv u + v \bmod (J, |x| \lneqq \mathbf{t})$, and since s is a primitive element of genus $\mathbf{t}$ in the separable group J, it follows from Lemma 5.2 that $m(u + v) / m(s)$.

$g(\mathbf{t}) / m(v_i)$ implies $g(\mathbf{t}) / m(v)$.

$g(\mathbf{t})_\infty = m(s)_\infty$, since $\mathbf{t}$ is the genus of s and $g(\mathbf{s})$ satisfies (b2).

$g(\mathbf{t})_f / g(|u_i|) / m(u_i)$, since $|u_i| < \mathbf{t}$ and $g(\mathbf{s})$ satisfies (c) and (b3). Therefore $g(\mathbf{t})_f / m(u)$. Thus it has been proved that $g(\mathbf{t}) / m(u + v) / m(s)$ and therefore that $g(\mathbf{t}) / g(S < J, \mathbf{t})$. Since $S(\mathbf{t}) \leqq S$, it follows that $g(S, \mathbf{t}) / g(S(\mathbf{t}), \mathbf{t}) = g(\mathbf{t})$, i.e., $g(\mathbf{s}) = g(S, \mathbf{s})$.

COROLLARY 6.7. *Suppose that $b \neq 0$ is an element of the separable group J and that S is the smallest regular group of J which contains b. Then*

$$g(S, \mathbf{s}) = \infty, \text{ if } m(b, \mathbf{s}+) = \infty \text{ or if } (J, \mathbf{s} \leqq |x|) = (J, \mathbf{s} < |x|);$$

If $m(b, \mathbf{s}+) \neq \infty$ and $(j, \mathbf{s} \leqq |x|) \neq (J, \mathbf{s} < |x|)$, then

$$g(S, \mathbf{s}) = m(b < J, \mathbf{s}+), \text{ if } m(b < J, \mathbf{s}) = \infty,$$

and if $m(b < J, \mathbf{s}) \neq \infty$, then $g(S, \mathbf{s})_f = m(b < J, \mathbf{s}+)_f$ and $g(S, \mathbf{s})_\infty$ is the infinite part of the numbers of genus $\mathbf{s}$.

Proof. As a consequence of Theorems 6.6 and 6.5c there exists one and only one regular subgroup S of J such that $g(S, \mathbf{s})$ has the value given in the Corollary 6.7. This regular subgroup S contains the element b, since the elements b_i of a primitive set with sum b are by Corollary 5.7 and by Theorem 6.5a contained in S. If finally T is a regular subgroup of J which contains b, then $S \leqq T$ by Lemma 6.2 and Theorem 6.5a, i.e., S is the smallest regular subgroup of J which contains b.

COROLLARY 6.8. *The elements b and b' of the separable group J are contained in the same regular subgroups of J if, and only if, $m(b, \mathbf{s}) = m(b', \mathbf{s})$ and $m(b, \mathbf{s}+) = m(b', \mathbf{s}+)$ for every genus $\mathbf{s}$.*

For the smallest regular subgroup of J which contains b depends by Corollary 6.7 only on the values of $m(b, \mathbf{s})$ and $m(b, \mathbf{s}+)$.

Thus the type of an element is generally not completely determined by the regular subgroups which contain the element.

COROLLARY 6.9. *If J is a separable group, the following statements are equivalent:*

(a) *Two elements of J are isotype in J if, and only if, they are contained in the same regular subgroups of J.*

(b) *There does not exist a genus $\mathbf{w}$ with the following properties:*

(b1) *$\mathbf{w}$ is a singular genus for J;*

(b2) *there exists a genus $\mathbf{s}$ such that $\mathbf{s} < \mathbf{w}$ and $(J, \mathbf{s} < |x|) \neq (J, \mathbf{s} \leqq |x|)$;*

(b3) *there exists an ordinary positive integer n which is relatively prime to the infinite part of the numbers of genus $\mathbf{w}$ such that $Z(J, n, \mathbf{w})$ contains elements of order n which are not conjugate.*

This is a consequence of Corollary 6.8, Lemma 5.8, Corollary 5.10b and Theorem 5.12.

NOTATION. *If $b \neq 0$ is an element of the separable group J, $\mathbf{s}$ a genus, then $q(b, \mathbf{s})'$ is either 1 or 2 and $q(b, \mathbf{s})' = 2$ if, and only if,*

(1) s *is singular for* J *and* 2 *relatively prime to the infinite part of the numbers of genus* s;

(2) $m(b, \mathsf{s}) \neq m(b, \mathsf{s}+)$ *and either*

(21) $m(b, \mathsf{s}) \neq \infty$, $q(b, \mathsf{s})$ *even or*

(22) $m(b, \mathsf{s}) = \infty$.

THEOREM 6.10. *Suppose that* $b \neq 0$ *is an element of the separable group* J, *that* C *is the smallest characteristic subgroup of* J *containing* b, *and that* R *is the greatest regular subgroup of* J *contained in* C.

(a) $C = R$ *if, and only if, at most one of the numbers* $q(b, \mathsf{s})' = 2$.

(b) *If* $R < C$, *then*

(b1) C/R *is a cyclic group of order* 2 (*and consists therefore of the elements* R *and* $R + b$);

(b2) $g(R < J, \mathsf{s}) = \infty$, *if* $(J, \mathsf{s} < |\, x \,|) = (J, \mathsf{s} \leqq |\, x \,|)$ *or if* $m(b, \mathsf{s}+) = \infty$, $g(R < J, \mathsf{s})_f = q(b, \mathsf{s})'m(b, \mathsf{s}+)_f$ *and*

$g(R < J, \mathsf{s})_\infty$ *is the infinite part of the numbers of genus* s, *if* $(J, \mathsf{s} < |\, x \,|) \neq (J, \mathsf{s} \leqq |\, x \,|)$ *and* $m(b, \mathsf{s}+) \neq \infty$.

Proof. Since $q(b, \mathsf{s})' \,/\, q(b, \mathsf{s})$, if $m(b, \mathsf{s}) \neq \infty$, there exists by Theorem 6.6 one and by Theorem 6.5c only one regular group R such that the values of $g(R, \mathsf{s})$ are those given in (b2). By Lemma 6.2, $R \leqq C$. $2b$ is an element of R by Lemma 5.8 and Theorem 6.5a.

If every $q(b, \mathsf{s})' = 1$, then R is by Corollary 6.7 the smallest regular subgroup of J containing b, and therefore $R = C$ is a regular subgroup of J.

If at least one of the $q(b, \mathsf{s})' = 2$, then b is not an element of R, as follows from Lemma 6.2 and Lemma 5.7, 5.8. But $2b$ is an element of R and the elements of R and $R + b$ form therefore a group C' which contains b and satisfies $R \leqq C' \leqq C$. C'/R is a cyclic group of order 2.

Let $b_1, \cdots, b_k$ be a primitive set with sum b and satisfying the conditions of Lemma 5.8. If s is the genus of b_i, then b_i is an element of R if, and only if, $q(b, \mathsf{s})' = 1$. If $q(b, |\, b_i \,|)' = 2$ for $1 \leqq i \leqq h$, then $1 \leqq h \leqq k$ and

$$b \equiv b' = \sum_{i=1}^{h} b_i \mod R.$$

Let α be a proper automorphism of J. Since $|\, b_i \,|$ is for $1 \leqq i \leqq h$ a singular genus, it follows from Lemma 5.6 that $b_i\alpha = r_i(b_i + b_i')$, where r_i is a rational number $\neq 0$ whose numerator and denominator are divisors of the infinite part of the numbers of genus $|\, b_i \,|$ and where b_i' is an element of $(J, |\, b_i \,| < |\, x \,|)$ satisfying $m(b_i) \,/\, m(b_i')$.

Now $\mathsf{s} < \mathsf{t}$ implies that $q(b, \mathsf{t})'m(b, \mathsf{t}+)_f \,/\, m(b, \mathsf{s}+)_f$ (if $q(b, \mathsf{t})'$ is defined), since the 2-value of $m(b, \mathsf{t}+)$ is smaller than the 2-value of $m(b, \mathsf{t})$ and therefore than the 2-value of $m(b, \mathsf{s}+)$, if $q(b, \mathsf{t})' = 2$. Furthermore, b_i' is the sum of a primitive set $b_{i1}, \cdots$ and, since $|\, b_i \,| < |\, b_{ij} \,|$, it follows that

$$q(b, |\, b_{ij} \,|)'m(b, |\, b_{ij} \,| +)_f \,/\, m(b, |\, b_i \,| +)_f \,/\, m(b_i) \,/\, m(b_i') \,/\, m(b_{ij}),$$

i.e., b_{ij} and consequently b_i' is an element of R.

If $r_i = r'_i r''^{-1}_i$, then r', r'' are both odd (if relatively prime) and since $2b_i$ and (by Corollary 5.3) $2r''^{-1}_i b_i$ are elements of R, it follows that $b_i - b_i\alpha \equiv 0 \bmod R$. Since R is a characteristic subgroup, C' is therefore also a characteristic subgroup, and since b is an element of $C' \leqq C$, it follows that $C' = C$.

If exactly one $q(b, \mathbf{s})' = 2$, i.e., $h = 1$, then b is the sum of a primitive set in J which is contained in C. C is therefore generated by the primitive elements in J which are contained in C, i.e., C is by Theorem 6.6 (and its proof) regular.

If $2 \leqq h$ (in the above notation), then b_1 and b_2 are not contained in R and therefore not in C, which is singular by (6.2.1).

COROLLARY 6.11. *The elements b and b' of the separable group J are contained in the same characteristic subgroups of J if, and only if, $m(b, \mathbf{s}) = m(b', \mathbf{s})$ and $m(b, \mathbf{s}+) = m(b', \mathbf{s}+)$ for every genus $\mathbf{s}$, i.e., if, and only if, b and b' are contained in the same regular subgroups of J.*

Proof. If b and b' are contained in the same characteristic subgroups of J, then they are contained in the same regular subgroups of J. If they are contained in the same regular subgroups, then $m(b, \mathbf{s}) = m(b', \mathbf{s})$ and $m(b, \mathbf{s}+) = m(b', \mathbf{s}+)$ for every genus $\mathbf{s}$ by Corollary 6.8. If b and b' satisfy these identities, then it follows from Theorem 6.10, since $q(b, \mathbf{s})' = q(b', \mathbf{s})'$, that $R(b) = R(b')$, where $R(u)$ is the greatest regular subgroup of the smallest characteristic subgroup $C(u)$ of J which contains u. If furthermore $b_1, \cdots, b_k$ is a primitive set with sum b and $b'_1, \cdots, b'_{k'}$ a primitive set with sum b', b_i not contained in $R(b)$, then $2b_i$ is contained in $R(b)$, $q(b, |b_i|_i)' = 2$ and therefore $q(b', |b_i|)' = 2$, i.e., $|b'_i| = |b_i|$ and $b'_i - b_i \equiv 0 \bmod R$, i.e., $C(b) = C(b')$. Note that $C(b)$ is completely determined by $R(b)$ and the set of genera $\mathbf{s}$ such that $q(b, \mathbf{s})' = 2$.

COROLLARY 6.12. *There exist singular subgroups of the separable group J if, and only if, there exist at least two different singular genera $\mathbf{s}$ such that the 2-values of the numbers of genus $\mathbf{s}$ are finite.*

This is a consequence of Theorem 6.10, since the join of regular subgroups is a regular subgroup, and since every characteristic subgroup S is the join of the groups $C(s)$ with s in S.

Survey of the singular subgroups. If C is any characteristic subgroup of the separable group J, then C contains a greatest regular subgroup $R(C)$ and is contained in a smallest regular subgroup $S(C)$. By Lemma 6.2 and Theorem 6.10 all the elements $\neq 0$ of $S(C)/C$ and $C/R(C)$ are of order 2.

Suppose now that C is singular and $b^* \neq 0$ an element of $C^* = C/R(C)$. Then denote by $\mathbf{S}(b^*)$ *the set of all the genera* $\mathbf{s}$ *such that* $q(b, \mathbf{s})' = 2$ *for every element b in the class b^**. $\mathbf{S}(b^*)$ contains at least two different genera and there exists in b^* an element b_0 such that $m(b_0, \mathbf{s}) \neq m(b_0, \mathbf{s}+)$ if, and only if, $\mathbf{s}$ is contained in $\mathbf{S}(b^*)$ and $q(b_0, \mathbf{s})' = 2$, if $\mathbf{s}$ is contained in $\mathbf{S}(b^*)$. $\mathbf{S}(b^* + c^*)$ contains exactly those genera which are contained either in $\mathbf{S}(b^*)$ or in $\mathbf{S}(c^*)$ but not in both these sets.

Thus in the set $\mathbf{T} = \mathbf{T}(C)$ of sets $\mathbf{S}(b^*)$ with b^* in C^* an addition has been defined, such that $\mathbf{T}$ becomes an abelian group ($\mathbf{S}(0) = 0$) and such that the correspondence $\mathbf{S}(b^*)$ between C^* and $\mathbf{T}$ is an isomorphism.

Finally if $\mathbf{s}$ is an element of $\mathbf{S}(b^*)$, then $\mathbf{s}$ is a singular genus for J and the 2-value of $g(R(C), \mathbf{s})$ is positive and finite.

If conversely C^* is an abelian group whose elements $\neq 0$ are of order 2, R a regular subgroup of the separable group J and $\mathbf{S}(b^*)$ a function of the elements of C^* which has all the mentioned properties, then there exists a characteristic subgroup C of J such that $R = R(C)$, $C^* = C/R(C)$ and $\mathbf{S}(b^*)$ is the above isomorphism between C^* and $\mathbf{T}(C)$.

Finally C is uniquely determined by $R(C)$ and $\mathbf{S}(b^*)$.

Characteristic decompositions. A direct decomposition of a group J is a characteristic decomposition, if all the summands are characteristic subgroups. Every characteristic direct summand is regular.[13]

The group J is characteristic irreducible, if J is a summand of every characteristic decomposition. It may happen that J is characteristic irreducible and that there exists a characteristic direct summand $\neq 0$, $\neq J$ of J. There exists at most one characteristic decomposition in characteristic irreducible direct summands.[13]

Suppose now that the group J is separable and partially reducible. The set $\mathbf{P}(J)$ of all the genera $\mathbf{s}$ with $(J, \mathbf{s} \leqq |x|) \neq (J, \mathbf{s} < |x|)$ can be decomposed into sets $\mathbf{S}$ such that

every genus is contained in exactly one set $\mathbf{S}$;

every $\mathbf{S}$ contains genera;

if $\mathbf{s} < \mathbf{t}$ are genera in $\mathbf{P}(J)$, then they are contained in the same $\mathbf{S}$;

every $\mathbf{S}$ is as small as possible.

This decomposition of $\mathbf{P}(J)$ is unique. Let now $J(\mathbf{S})$ be the join of the subgroups $(J, \mathbf{s} \leqq |x|)$ with $\mathbf{s}$ in $\mathbf{S}$. Then J is the direct sum of the groups $J(\mathbf{S})$ and this decomposition is a characteristic decomposition into characteristic irreducible, direct summands.

7. **Types of subgroups.** The following concepts and facts concerning *abelian groups without elements of infinite order* will be used in the course of this section.

If F is an abelian group without elements of infinite order, then the *order* of F is the l.c.m. of the orders of its finite subgroups; the *rank* $m(F)$ of F is the smallest number m such that every finite subgroup of F can be decomposed into a direct sum of at most m cyclic groups. (Only groups of finite rank will be considered.)

If the rank $m(F)$ of the abelian group F without elements of infinite order is finite, there exists a *canonical decomposition* of F into a direct sum of $m(F)$ groups F_i of rank 1 such that the order of F_i is a divisor of the order of F_{i+1}. The orders of the summands F_i of a canonical decomposition do not depend on the particular canonical decomposition and determine the structure of F completely. They are the *invariants* $m_1(F), \cdots, m_{m(F)}(F)$ of F.

If F is in particular a finite group, then the groups F_i are cyclic groups and

[13] See R. Baer, *The decomposition of abelian groups into direct summands*, Quart. Jour. of Math., (2), vol. 6 (1935), pp. 222–232.

the subset B of F is a *basis* of F, if the cyclic subgroups of F, generated by the elements in B, form a canonical decomposition of F.

If F is of finite rank and S a subgroup of F, then S and F/S are of finite rank.

If A is any abelian group, the subset $F(A)$ of all the elements of finite order in A is a subgroup of A and $A/F(A)$ does not contain elements $\neq 0$ of finite order.

If J is an abelian group without elements $\neq 0$ of finite order, S a subgroup of J, then the closed subgroup $\bar{S}$ of J which is generated by S satisfies $F(J/S) = \bar{S}/S$. If S is of finite rank, then $F(J/S)$ is of finite rank and $m(F(J/S)) \leqq r(S)$.

If S is a subgroup of the (abelian) group J (without elements $\neq 0$ of finite order), then the *defect* of an element $b \neq 0$ of S in J is

$$d(b < S < J) = m(b < J) \dot{:} m(b < S).$$

An element $b \neq 0$ of S has the same genus in S and in J if, and only if, its defect is an ordinary integer.

If $R' \neq 0$ is a subgroup of the rational group R, all the elements $\neq 0$ of R' have the same defect $(R:R')$. Then $(R:R')$ is the order of the group R/R' of rank 1 and $R = (R:R')^{-1}R'$.

Lemma 7.1. (a) *The completely reducible group J of finite rank is finite* mod *its subgroup S if, and only if,*

(a′) $r(S) = r(J)$,

(a″) *the elements $\neq 0$ of S have the same genus in S and in J.*

(b) *If S is a subgroup of the direct sum J of a finite number of isomorphic rational groups (all of genus* **s**), *and if $F = J/S$ is finite, there exist decompositions*:

$$\text{(b+)} \qquad J = \sum_{i=1}^{r(J)} R_i, \qquad S = \sum_{i=1}^{m(F)} m_i(F)R_i + S', \qquad S' = \sum_{i=m(F)+1}^{r(J)} R_i,$$

where the R_i are groups of genus **s**, *S' is a greatest closed subgroup of J contained in S, $m_i(F)R_i$ is for $1 \leqq i \leqq m(F)$ the intersection of S and R_i, $m(F)$ the finite rank and $m_i(F)$ the finite invariants of the finite group $F = J/S$, and $m(F) \leqq r(J)$.*

Proof. If J is a completely reducible group of finite rank and J/S is finite, then $r(J) = r(S)$, since every element of J has finite order mod S. If $b \neq 0$ is an element of S, $\bar{b}$ the closed subgroup of J, generated by b, $\bar{b}'$ the intersection of S and $\bar{b}$, then $\bar{b}/\bar{b}'$ represents exactly a subgroup of J/S, i.e., $(\bar{b}:\bar{b}')$ is finite and b has the same genus in S and in J.

If, conversely, (a′) and (a″) are satisfied, J is the direct sum of the rational groups J_i, J_i' the intersection of S and J_i, S' the direct sum of the groups J_i', then $S' \leqq S$, $(J_i:J_i')$ finite, i.e., J/S' is finite and therefore $J/S = (J/S')/(S/S')$ is finite.

Suppose now that J is a direct sum of a finite number of isomorphic rational groups of genus **s** and that $J/S = F$ is finite. There exists a greatest closed subgroup S' of J such that $S' \leqq S$. S' is by Corollary 3.5 a direct summand of J, since $r(J)$ is finite and S' is closed. If $J = S' + J''$, then $S = S' + S''$, where S'' is the intersection of S and J''. $F = J/S$ and J''/S'' are isomorphic finite groups and there exists therefore a basis B of J'' mod S''.

(7.1.1) Every basis B of J'' mod S'' is a greatest independent subset of J''.

For if $\sum_{b \text{ in } B} c_b b = 0$ with integer, relatively prime coefficients c_b, then $c_b b \equiv 0$ mod S'', and therefore $c_b \equiv 0$ mod $m_1(F)$, i.e., $c_b = 0$ and B is independent.

In order to prove that B contains $r(J'')$ elements, denote by $\bar{B}$ the closed subgroup of J'' generated by B. Since every class of J''/S'' contains elements of $\bar{B}$, there exists corresponding to every element $u \neq 0$ of J'' an element $s(u)$ such that $s(u) \equiv 0$ mod S'', $u \equiv s(u)$ mod $\bar{B}$. Since 0 is the only closed subgroup of J'' contained in S'', there exists corresponding to every element $s \neq 0$ of S'' a positive integer $w(s)$ such that $w(s)^{-1}s$ exists in J, but not in S''. There exists, therefore, corresponding to every element v of J'' a chain s_i, t_i satisfying

$$\begin{array}{llll} v = t_0, & s_i = 0, \quad \text{if } t_i = 0; & t_{i+1} = 0, & \text{if } s_i = 0, \\ & s_i = s(t_i), \text{ if } t_i \neq 0; & t_{i+1} = w(s_i)^{-1}s_i, & \text{if } s_i \neq 0. \end{array}$$

Then $v \equiv \prod_{i=0}^{k-1} w(s_i)\, s_k$ mod $\bar{B}$, if $s_i \neq 0$ for $i < k$, and therefore $v \equiv 0$ mod $\bar{B}$, if at least one $s_k = 0$.

If none of the $s_i = 0$, none of the numbers $w(s_i)$ is relatively prime to the finite order f of the finite group F and there exists therefore a prime number $p \mid f$ (p therefore relatively prime to the infinite parts of the multiplicities of elements in J'') such that

$$v \equiv p^i z_i \bmod \bar{B}$$

has for every i a solution z_i in J''. Since $\bar{B}$ is by Corollary 3.5 a direct summand of J'', this implies that $v \equiv 0$ mod $\bar{B}$, i.e., $J'' = \bar{B}$ and this completes the proof of (7.1.1).

Let now $b_1, \cdots, b_m$ form a basis of J'' mod S'' such that $m = m(F)$ and $m_i(F)$ is the order of b_i mod S''. Denote by B_i the closed subgroup of J generated by $b_1, \cdots, b_i$. Then B_1 is a rational group, $B_m = J''$ by (7.1.1), and $B_{i+1} = B_i + R_{i+1}$ by Corollary 3.5. Then $b_{i+1} = u_{i+1} + v_{i+1}$ with u_{i+1} in B_i, v_{i+1} in R_{i+1} and the elements $v_1 = b_1, v_2, \cdots, v_m$ form a basis of J'' which is also a basis of J'' mod S''. This shows that the groups R_i together with S' define a decomposition of J which meets the requirements of (b+).

Remark. The following example shows that it is impossible to omit in Lemma 7.1b the assumption that all the elements $\neq 0$ of J have the same genus in J.

A basis of the group J is formed by the elements b', b'' with $m(b' < J) = 1$, $m(b'' < J) = 2^\infty$ and the subgroup S is generated by the elements $3b' + b''$, $2^{-i}3b''$. Then J/S is a cyclic group of order 3, but there does not exist a decomposition (b+) of J, since a multiple of b'' is contained in every basis of J.

DEFINITION 7.2. *The direct decomposition*

$$J = \sum_v J_v + J'$$

of the group J is a complete reduction of the subgroup S in J, if the groups J_v are rational and $S = \sum_v S_v$ for S_v = intersection of S and J_v.

If there exists a complete reduction of the subgroup S in J, then S is *completely reducible in* J.

If the subgroup S of J is completely reducible in J, then S is completely reducible. The converse is not true, since e.g., rational subgroups S of separable groups J are completely reducible in them if, and only if, all the elements $\neq 0$ in S are primitive in J. It is a consequence of Lemma 7.1b that S is completely reducible in J, if J/S is finite, $r(J) = r(S)$, J is completely reducible and all the elements $\neq 0$ in J have the same genus in J.

COROLLARY 7.3. *Suppose that J and its subgroup S have the same finite rank.*

(a) *S is completely reducible in J, all the elements $\neq 0$ of J have the same genus $\mathbf{g}$ in J and all the elements $\neq 0$ of S have the same genus $\mathbf{s}$ ($\leqq \mathbf{g}$) in S if, and only if, there exists a subgroup T of J such that*

(1) $S \leqq T \leqq J$,

(2) *all the elements $t \neq 0$ of T have the same defect $d = d(t < T < J)$ in J (and therefore $J = d^{-1}T$),*

(3) *T is a direct sum of isomorphic rational groups,*

(4) *all the elements $\neq 0$ of S have the same genus in S and in T.*

(b) *If the subgroup S of J satisfies the conditions of* (a), *then there exist decompositions into rational groups:*

$$(\mathrm{b}+) \qquad J = \sum_{i=1}^{r(J)} J_i, \qquad S = \sum_{i=1}^{r(J)} S_i,$$

where S_i = intersection of S and J_i, $|J_i| = \mathbf{g}$, $|S_i| = \mathbf{s}$, $(J_i:S_i) = m_i(J/S)$ = i-th invariant of the group J/S without elements of infinite order for $1 \leqq i \leqq m(J/S)$.

Note that either all the numbers $(J_i:S_i)$ are ordinary integers—the case dealt with in Lemma 7.1—or all the $(J_i:S_i)$ have the same genus and $m(J/S) = r(J) = r(S)$.

Proof. Suppose first that S is completely reducible in J, that all the elements $\neq 0$ of J have the same genus $\mathbf{g}$, and that all the elements $\neq 0$ of S have the same genus $\mathbf{s}$ in S. Then there exist decompositions into rational groups

$$J = \sum_{i=1}^{r(J)} R_i, \qquad S = \sum_{i=1}^{r(J)} R'_i, \qquad |R_i| = \mathbf{g}, \qquad |R'_i| = \mathbf{s},$$

where R'_i is the intersection of S and R_i.

Denote by d the g.c.d. of all the $d(s < S < J)$ with $s \neq 0$ in S and by d^* the g.c.d. of the numbers $(R_i:R'_i)$. If $s \neq 0$ is an element of S, then $s = \sum_{i=1}^{r(J)} s_i$ with s_i in R'_i. Then $m(s < S)$ is the g.c.d. of the numbers $m(s_i < R'_i)$ and $m(s < J)$ is the g.c.d. of the numbers $m(s_i < R_i)$. If $s_i \neq 0$, there exists a uniquely determined ordinary positive integer h_i which is relatively prime to the infinite part of the numbers of genus $\mathbf{g}$ such that

$$m(s_i < J) = h_i m(s < J)$$

and a uniquely determined ordinary positive integer k_i which is relatively prime to the infinite part of the numbers of genus **s** such that

$$m(s_i < S) = k_i m(s < S).$$

Then $h_i d(s < S < J) = k_i(R_i:R_i')$, and since the numbers h_i with $s_i \neq 0$ are relatively prime, and also the numbers k_i with $s_i \neq 0$ are relatively prime, it follows that $d(s < S < J)$ is the g.c.d. of the numbers $(R_i:R_i')$ with $s_i \neq 0$, i.e., that

$$d = d^*.$$

Since $r(J)$ is finite and all the numbers $(R_i:R_i')$ have the same genus, it follows that every $n_i = (R_i:R_i') \div d$ is an ordinary integer. Denote now by R_i'' the subgroup of R_i which satisfies $n_i R_i'' = R_i'$. Then $(R_i:R_i'') = d$, $(R_i'':R_i') = n_i$ and the direct sum

$$T = \sum_{i=1}^{r(J)} R_i''$$

satisfies (1)–(4).

Suppose secondly that there exists a subgroup T of J which satisfies (1)–(4). Then there exists by Lemma 7.1 a decomposition

$$T = \sum_{i=1}^{r(J)} T_i, \qquad S = \sum_{i=1}^{r(J)} S_i$$

such that S_i is the intersection of S and T_i, $|S_i| = |T_j| = \mathbf{s}$ and the numbers $(T_i:S_i)$ with $1 \leqq i \leqq m(T/S)$ are the invariants of the finite group T/S.

If $J_i = d^{-1}T_i$ is the closed subgroup of J generated by T_i, then $J = \sum_{i=1}^{r(J)} J_i$, $|J_i| = \mathbf{t}$ and $(J_i:S_i) = d(T_i:S_i)$, i.e., these decompositions of J and S are decompositions (b+).

If finally there exist decompositions (b+) of J and S, then S is completely reducible in J, all the elements $\neq 0$ of J have the same genus in J and all the elements $\neq 0$ of S have the same genus in S.

If J and its subgroup S have the same finite rank; if J is completely reducible; if all the elements $\neq 0$ of J have the same genus in J; and if all the elements $\neq 0$ of S have the same genus in S, then it can happen that S is *not* completely reducible in J. This is shown by

Example 7.4. Let J be a group with a basis b', b'' such that $m(b' < J) = m(b'' < J) = p^\infty$ for a given prime number p. Choose the ordinary integers c_i in such a way that $c = \sum_{i=0}^{\infty} c_i p^i$ is an irrational p-adic number and $0 \leqq c_i < p$.

Denote by S the subgroup of J, generated by the elements b', b'', $p^{-i}(b' + \sum_{j=0}^{i-1} c_j p^j b'')$ for $i = 1, 2, \cdots$. Then all the multiplicities $m(s < S)$ of elements

$s \neq 0$ in S are ordinary integers, S/pS is a cyclic group of order p and S is therefore not even completely reducible (in S).[14]

Corollary 7.5. *Suppose that the group J and its subgroup S satisfy:*

(1) *S and J have the same finite rank;*

(2) *J is a direct sum of isomorphic rational groups (of genus $\mathbf{g}$);*

(3) *The* g.c.d. *$d(S < J)$ of the defects $d(s < S < J)$ of the elements $s \neq 0$ of S in J has the same genus as these defects;*[15]

(4) $d(S < J)_\infty = 1$.

Then S is completely reducible in J and all the elements $\neq 0$ of S have the same genus in S.

Proof. There exists by (3) to every element $s \neq 0$ of S an ordinary integer $h(s)$ such that $d(s < S < J) = h(s)d(S < J)$. Therefore $m(s < J)d(S < J) = h(s)m(s < S)$ by (4) and all the elements $\neq 0$ of S have consequently by (2) the same genus $\mathbf{s}$ in S.

If U is the subgroup of J which satisfies $U = d(S < J)^{-1}S$ and $S = d(S < J)U$ by (4), then all the elements $\neq 0$ of U have the same genus in U and in J. There exists therefore by Lemma 7.1 decompositions

$$J = \sum_{i=1}^{r(J)} J_i, \qquad U = \sum_{i=1}^{r(J)} U_i,$$

U_i = intersection of U and J_i. Thus $S = d(S < J)U = \sum_{i=1}^{r(J)} d(S < J)U_i$ and $d(S < J)U_i$ = intersection of S and J_i, i.e., S is completely reducible in J.

Notation. *If S is a subset of the group J, then $\bar{S}$ is the closed subgroup of J, generated by S.*

If $\mathbf{s} \leqq \mathbf{t}$ are genera and S a subgroup of the group J, then

$(S < J, \mathbf{t}, \mathbf{s})$ = *intersection of* $(J, \mathbf{t} \leqq | x < J |)$ *and* $(S, \mathbf{s} \leqq | x < S |)$,

$(S < J, \mathbf{t}, \mathbf{s}+)$ = *intersection of* $(J, \mathbf{t} \leqq | x < J |)$ *and* $(S, \mathbf{s} < | x < S |)$,

$(S < J, \mathbf{t}, \mathbf{s})^*$ = *join of* $(S < J, \mathbf{t}, \mathbf{s})$ *and* $\overline{(S < J, \mathbf{t}, \mathbf{s}+)}$,

$n(S < J, \mathbf{t}, \mathbf{s})$ = $r(\overline{(S < J, \mathbf{t}, \mathbf{s})/(S < J, \mathbf{t}, \mathbf{s}+)})$,

$N(S < J, \mathbf{t}, \mathbf{s})$ = $\overline{(S < J, \mathbf{t}, \mathbf{s})}/(S < J, \mathbf{t}, \mathbf{s})^*$.

These notations and Corollary 7.3 imply the following statement.

(7.6) *If the formulas*

$$J = \sum_v \bar{S}_v + J', \qquad S = \sum_v S_v,$$

where S_v is both a rational group and the intersection of S and $\bar{S}_v$, give a complete reduction of S in J, and if for $\mathbf{s} \leqq \mathbf{t}$

[14] The group S belongs to a class of groups which has been discussed recently by A. Kurosch, Ann. of Math., vol. 38 (1937), pp. 175–204.

[15] Note that the group J and its subgroup S which have been considered as Example 7.4 satisfy (1)–(3).

$S(\mathbf{t}, \mathbf{s})$ is the direct sum of those groups S_v which satisfy $\mathbf{s} = | S_v |$ and $\mathbf{t} = | \bar{S}_v |$, then

$$(S < J, \mathbf{t}, \mathbf{s}) = \sum_{\mathbf{t} \leqq \mathbf{v}} \sum_{\mathbf{s} \leqq \mathbf{u}} S(\mathbf{v}, \mathbf{u}), \qquad (S < J, \mathbf{t}, \mathbf{s}+) = \sum_{\mathbf{t} \leqq \mathbf{v}} \sum_{\mathbf{s} < \mathbf{u}} S(\mathbf{v}, \mathbf{u}),$$

$$n(S < J, \mathbf{t}, \mathbf{s}) = r(S(\mathbf{t}, \mathbf{s})) = r(\overline{S(\mathbf{t}, \mathbf{s})}), \qquad N(S < J, \mathbf{t}, \mathbf{s}) = \overline{S(\mathbf{t}, \mathbf{s})}/S(\mathbf{t}, \mathbf{s}).$$

If the ranks $n(S < J, \mathbf{t}, \mathbf{s})$ are finite, there exists a decomposition

$$J = \sum_v \bar{R}_v + J', \qquad S = \sum_v R_v$$

such that the rational group R_v is the intersection of S and $\bar{R}_v$, $S(\mathbf{t}, \mathbf{s})$ is the direct sum of the groups R_v with $\mathbf{s} = | R_v |$, $\mathbf{t} = | \bar{R}_v |$, and the numbers $(\bar{R}_v : R_v) \neq 1$ with $\mathbf{s} = | R_v |$, $\mathbf{t} = | \bar{R}_v |$ are exactly the invariants of $N(S < J, \mathbf{t}, \mathbf{s})$.

COROLLARY 7.7. *Suppose that S is a subgroup of finite rank of the separable group J. Then S is completely reducible in J if, and only if, there exists a direct decomposition $S = \sum_{\mathbf{t}} \sum_i S_{\mathbf{t}i}$ such that*

(a) *$S_{\mathbf{t}i}$ is completely reducible in $\bar{S}_{\mathbf{t}i}$;*

(b) *every element $\neq 0$ of $S_{\mathbf{t}} = \sum_i S_{\mathbf{t}i}$ is a primitive element of genus $\mathbf{t}$ in J;*

(c) $\bar{S}_{\mathbf{t}} = \sum_i \bar{S}_{\mathbf{t}i}$.

Proof. If S is completely reducible in J, then $\bar{S}$ is a completely reducible direct summand of J and the groups $S(\mathbf{t}, \mathbf{s})$ of (7.6) show the necessity of the condition. If conversely the condition is satisfied, then every $\bar{S}_{\mathbf{t}}$ is by (b) and Theorem 4.2 a direct summand of J. $\bar{S}$ is therefore the direct sum of the groups $\bar{S}_{\mathbf{t}}$ and by Theorem 4.2 a direct summand of J. S is completely reducible in J as a consequence of this fact and of (a) and (c).

THEOREM 7.8. *Suppose that the subgroups S and T are both completely reducible in the group J, and that all the ranks $n(S < J, \mathbf{t}, \mathbf{s})$ are finite. Then S and T are isotype in J if, and only if,*

(a) *$n(S < J, \mathbf{t}, \mathbf{s}) = n(T < J, \mathbf{t}, \mathbf{s})$ for any two genera $\mathbf{s} \leqq \mathbf{t}$;*

(b) *$N(S < J, \mathbf{t}, \mathbf{s})$ and $N(T < J, \mathbf{t}, \mathbf{s})$ are isomorphic for any two genera $\mathbf{s} \leqq \mathbf{t}$;*

(c) *J/S and J/T are isomorphic.*

Proof. The necessity of the conditions is obvious. If the conditions (a) to (c) are satisfied, there exist by (7.6) direct decompositions

$$J = \sum_v \bar{S}_v + J' = \sum_v \bar{T}_v + J'', \qquad S = \sum_v S_v, \qquad T = \sum_v T_v$$

such that the rational groups S_v are the intersections of S and $\bar{S}_v$, the rational groups T_v the intersections of T and $\bar{T}_v$, the numbers $(\bar{S}_v : S_v) \neq 1$ are the invariants of $N(S < J, | \bar{S}_v |, | S_v |)$, the numbers $(\bar{T}_v : T_v) \neq 1$ are the invariants of $N(T < J, | \bar{T}_v |, | T_v |)$. (Note that by (a) also the ranks $n(T < J, \mathbf{t}, \mathbf{s})$ are finite.)

By (a), (b) and (7.6) it is possible to denote the groups S_v in such a way that

$$| \bar{S}_v | = | \bar{T}_v |, \qquad | S_v | = | T_v |, \qquad (\bar{S}_v : S_v) = (\bar{T}_v : T_v)$$

for every v. Since J' and $(J/S)/F(J/S)$, J'' and $(J/T)/F(J/T)$ are isomorphic, (c) implies that J' and J'' are isomorphic. Therefore there exists a proper automorphism of J which maps $\bar{S}_v$ upon $\bar{T}_v$, J' upon J'', and since this automorphism also maps S_v upon T_v, i.e., S upon T, S and T are isotype in J.

COROLLARY 7.9. *Suppose that S and T are subgroups of finite rank of the separable group J, and that S and T are both completely reducible in J. Then S and T are isotype in J if, and only if, for any two genera* $\mathbf{s} \leqq \mathbf{t}$

(a) $n(S < J, \mathbf{t}, \mathbf{s}) = n(T < J, \mathbf{t}, \mathbf{s})$;

(b) $N(S < J, \mathbf{t}, \mathbf{s})$ *and* $N(T < J, \mathbf{t}, \mathbf{s})$ *are isomorphic.*

Proof. Suppose that the conditions (a) and (b) are satisfied. By Corollary 4.3 J has a completely reducible direct summand D of finite rank which contains both S and T. Then it follows from (a), (b) and the finiteness of the ranks n that D/S and D/T are isomorphic, and S and T are therefore isotype in J by Theorem 7.8.

If the ranks $n(S < J, \mathbf{t}, \mathbf{s})$ are finite, the ranks n and the groups N satisfy

(i) $m(N(S < J, \mathbf{t}, \mathbf{s})) \leqq n(S < J, \mathbf{t}, \mathbf{s})$, and the equality holds, if $\mathbf{s} < \mathbf{t}$;

(ii) $\sum_{\mathbf{t}} n(S < J, \mathbf{t}, \mathbf{s}) \leqq r(J(\mathbf{t})^*)$;

(iii) to every invariant $m_i(N(S < J, \mathbf{t}, \mathbf{s}))$ there exists a number t of genus $\mathbf{t}$ and a number s of genus $\mathbf{s}$ such that $t \dot{:} s = m_i(N(S < J, \mathbf{t}, \mathbf{s}))$.

If $\sum_{\mathbf{t}, \mathbf{s}} n(S < J, \mathbf{t}, \mathbf{s})$ is finite and J is separable, the above conditions are also sufficient for the existence of a subgroup S of J with these given ranks n and groups N.

The following particular cases of the Corollary 7.9 may be mentioned. Suppose that J is a separable group, that S and T are subgroups of finite rank, that all the elements $\neq 0$ of S and of T are primitive elements of the given genus $\mathbf{g}$ in J, that either all the elements $\neq 0$ in S and in T have in S and T respectively the genus $\mathbf{g}$; or all the elements $\neq 0$ in S and in T have in S and T the genus $\mathbf{s}(< \mathbf{g})$, and that S and T are completely reducible in J. Then S and T are isotype in J if, and only if,

(a) $r(S) = r(T)$,

(b) $F(J/S)$ and $F(J/T)$ are isomorphic.

Since $r(J) = r(S) + r((J/S)/F(J/S)) = r(S) + r(J/\bar{S})$, we have also the following result. If the subgroups S and T of the separable group J satisfy the above assumptions, and if the rank of J is finite (and J therefore completely reducible), then S and T are isotype in J if, and only if, J/S and J/T are isomorphic.

If finally J is a direct sum of a finite number of infinite cyclic groups, and S is any subgroup of J, then J/S is finite and the type of S in J is completely determined by the structure of J/S.

The following remark may give an idea of what happens if subgroups are considered which are not completely reducible in the whole group. Let J be a complete group. Then the type of the subgroup S of J in J is completely determined by the structure of S and the structure of $J/\bar{S}$.

Chapter IV. Reducibility and separability

8. Decomposition into isomorphic rational groups.

LEMMA 8.1. *Suppose that S is a subgroup of the group J and that all the elements of the class $b^* \neq 0$ of J/S have the same genus* $\mathbf{s}$ *in J. Then*

(a) $S \leqq (J, \mathbf{s} \leqq | x |)$.

(b) $\mathbf{s}$ *is the genus of b^* in J/S if, and only if, there exists in b^* an element b such that $m(b < J) = m(b < J/S)$.*

Proof. If $s \neq 0$ is an element of S, b an element of b^*, then b and $b + s$ have both genus $\mathbf{s}$ and $\mathbf{s}$ is therefore a divisor of the genus of s.

If there exists in b^* an element b such that $m(b < J) = m(b < J/S)$, then b and b^* have the same multiplicity and therefore the same genus.

Suppose now that $\mathbf{s}$ is the genus of b^* and b' an element of J in b^*. Then $m(b' < J) / m(b' < J/S) = m(b^* < J/S)$ and there exists an ordinary positive integer h which is relatively prime to $m(b')_\infty$ such that $hm(b' < J) = m(b' < J/S)$. Denote by h' the greatest divisor of $m(b' < J/S)$ such that the same prime numbers divide h and h'. Then h' is an ordinary positive integer and there exists an element z in J such that $h'z \equiv b' \bmod S$. There exists, as above, an ordinary positive integer k which is relatively prime to $m(h'z)_\infty$ such that $km(h'z < J) = m(h'z < J/S) = m(b' < J/S) = m_0$. Since $m(h'z) = h'm(z)$, $m_0h'^{-1}$ exists and is relatively prime to h' (by its definition) and k and h are therefore relatively prime. There exist ordinary integers k'', h'' such that $kk'' + hh'' = 1$ and

$$b = h''hb' + k''kh'z \equiv b' \bmod S$$

satisfies $m(b < J) / m_0 / \text{g.c.d. } h''hm(b' < J), k''km(h'z < J) / m(b < J)$, i.e., $m(b < J) = m(b < J/S)$.

COROLLARY 8.2. *Suppose that $S < J$ and that all the elements of J which are not contained in S have the same genus* $\mathbf{s}$ *in J. Then*

(a) $(J, \mathbf{s} < | x |) \leqq S < J = (J, \mathbf{s} \leqq | x |)$.

(b) *All the elements $\neq 0$ of J/S have the genus* $\mathbf{s}$ *if, and only if, there exists corresponding to every element b of J an element b' of J such that*

$$b \equiv b' \bmod S, \qquad m(b' < J) = m(b' < J/S)(= m(b < J/S)).$$

This follows from Lemma 8.1 and the fact that every element of $(J, \mathbf{s} < | x |)$ is the sum of elements of S.

DEFINITION 8.3. *The class Γ_1 consists of all countable groups. If ν is a finite or infinite ordinal greater than 1, then Γ_ν consists of all groups J satisfying*

(a) *J does not belong to a class Γ_μ with $\mu < \nu$;*

(b) *There exists a closed subgroup S of finite rank of J such that J/S is a direct sum of groups which belong to the join Γ^ν of the classes Γ_μ with $\mu < \nu$.*

The additive group P of the integer p-adic numbers belongs to the second class and the direct sum of an infinity of such groups P belongs to the third class. The additive group of all the sequences of integers does not belong to any of these classes, as will be proved in the section on vector-groups.

THEOREM 8.4. *Assume that all the elements of the group J which are not con-*

tained in the subgroup S of J have the same genus **s**. *Then J is the direct sum of S and rational groups* (*of genus* **s**) *if, and only if,*

(a) *S is a closed subgroup of J;*

(b) *the elements $\neq 0$ of J/T have the genus* **s** *in J/T for every closed subgroup T of J such that $S \leqq T$ and $r(T/S)$ is finite;*

(c) *J/S belongs to a class Γ_ν.*

Proof. A. Suppose that $J = S + R$, where R is completely reducible. Then S is a closed subgroup of J. If $S \leqq T \leqq J$, then $T = S + T'$, where T' is the intersection of T and R. If T is a closed subgroup of J and $r(T/S)$ is finite, then T' is a closed subgroup of R and $r(T')$ is finite. T' is by Corollary 4.4 a direct summand of R, i.e., $R = T' + R'$ and consequently all the elements $\neq 0$ of J/T are of genus **s** in J/T, since J/T and R' are isomorphic. Since finally R and J/S are isomorphic, J/S belongs either to Γ_1 or to Γ_2.

B. Suppose now that the conditions (a) to (c) are satisfied by the subgroup S of J.

1. $r(J/S) = 1$.

Then there exists by Corollary 8.2 an element b in J such that $b \not\equiv 0 \bmod S$, $m(b < J) = m(b < J/S)$. The elements of the closed subgroup $\bar{b}$ of J, generated by b, represent therefore exactly the classes of J/S and consequently $J = S + \bar{b}$.

2. J/S is countable.

Then there exist closed subgroups T_i of J for $-1 \leqq i < r(J/S)$ such that $S = T_{-1}$, $T_{i-1} < T_i$, $r(T_i/T_{i-1}) = 1$, and such that J is the join of the groups T_i (if $r(J/S) = m + 1$ is finite, then $J = T_m$). Since $r(T_i/S) = i + 1$, it follows from 1 that $T_i = T_{i-1} + R_i$ for $-1 < i < r(J/S)$, where R_i is a rational group of genus **s**. Since J is the join of the groups T_i, it follows that

$$J = S + \sum_{0 \leqq i < r(J/S)} R_i .$$

3. J/S belongs to the class Γ_ν.

Since for $\nu = 1$ the theorem has been proved in 2, it can be assumed that the theorem holds true for subgroups S' of groups J' which satisfy (a), (b) and such that J'/S' belongs to Γ^ν.

Since J/S belongs to Γ_ν, there exists a closed subgroup W of J such that $S \leqq W$, $r(W/S)$ is finite and $J^* = J/W$ is the direct sum of groups J_v^*, belonging to Γ^ν. Let J_v be the subgroup of J which contains W and satisfies $J_v/W = J_v^*$. Then the subgroup W of J_v satisfies the conditions (a), (b), and since J_v^* belongs to Γ^ν, $J_v = W + W_v$, where W_v is completely reducible. Suppose now that b is an element of W, that the elements b_i belong to different groups W_v and that $b + \sum_i b_i = 0$. Then every b_i belongs to W, since J^* is the direct sum of the groups J_v^*, i.e., $b_i = 0$, since 0 is the intersection of W and W_v, i.e., $b = 0$. Thus the groups W, W_v are independent and, since every class of J/W contains elements of the sum of the groups W_v, it follows that

$$J = W + \sum_v W_v .$$

Since $r(W/S)$ is finite, W/S is countable and it follows from the case, treated in 2, that $W = S + V$, with V completely reducible, and consequently $J = S + R$, with $R = V + \sum_v W_v$ completely reducible.

Corollary 8.5. *The group J is a direct sum of isomorphic rational groups (of genus* **s**), *if, and only if,*

(a) *the elements $\neq 0$ of J/T all have the same genus* **s** *for every closed subgroup T of finite rank of J;*

(b) *J belongs to a class Γ_v.*

This follows from Theorem 8.4 by choosing $S = 0$.

Remarks. 1. If J is complete, then J is completely reducible and in this case the conditions (a) and (b) are not needed.

2. If J is a closed subgroup of the group P of all the integer p-adic numbers, then J/pJ is a group of order p. Since all the elements $\neq 0$ in S have the same genus, J is directly irreducible, and satisfies condition (b) and a weaker form of condition (a). This shows that condition (a) cannot be omitted, since there exist closed subgroups J of every finite rank in P.

3. The additive group of all the sequences of integers furnishes an example of a group, satisfying condition (a), but not condition (b), as will be proved in the section on vector-groups.

The following criterion provides a handy method for constructing a basis in some simple cases.

Corollary 8.6. *Suppose that $n \neq 1$ is the product of a finite number of different prime numbers, that* **n** *is the genus of those numbers whose infinite part is divisible by all prime numbers which are relatively prime to n and that J is a direct sum of a finite number of rational groups of genus* **n**. *Then the subset B of J is a basis of J if, and only if, B consists of multiples of the elements of a basis of J* mod nJ.

Proof. If first B is a basis of J, the elements $m(b < J)_f^{-1}b$ with b in B form a basis of J mod nJ. Assume conversely that the elements in B are multiples of the elements of the basis B' of J mod nJ. B' is an independent subset of J, since every relation between elements of B' implies a relation mod nJ. Denote now by $\bar{B}$ the closed subgroup of J, generated by B (or by B'). If w is any element in J, there exists an element w' in $\bar{B}$ such that $w \equiv w'$ mod nJ. The congruence $nz \equiv w$ mod $\bar{B}$ has therefore a solution z in J for every element w, i.e., $J/\bar{B} = n(J/\bar{B})$. Corollary 8.5 implies therefore that $J = \bar{B}$, i.e., that B' and B are greatest independent subsets of J. Consequently there exist corresponding to every element $w \neq 0$ in J relatively prime integers $k \neq 0$, $k(b)$ such that

$$kw = \sum_{b \text{ in } B'} k(b)b.$$

k and n are relatively prime, since B' is a basis of J mod nJ and therefore a common divisor of k and n divides every $k(b)$. Hence every element $k^{-1}k(b)b$ exists in J, i.e., B' and therefore B is a basis of J.

Note that the finiteness of $r(J)$ is a consequence of the validity of the above criterion and of the other assumptions of Corollary 8.6.

COROLLARY 8.7. *Suppose that all the elements of the group J which are not contained in the subgroup S of J have the same genus* $\mathbf{s}$ *in J. Then the following three statements are equivalent.*

(a) *J is a direct sum of S and a completely reducible group.*

(b) *J/S is a direct sum of rational groups of genus* $\mathbf{s}$.

(c) *J/S is separable, belongs to a class Γ_ν and all the elements $\neq 0$ in J/S have the genus* $\mathbf{s}$.

Proof. If J is the direct sum of S and the completely reducible group R, then R is a direct sum of groups of genus $\mathbf{s}$ and R and J/S are isomorphic, i.e., J/S is a direct sum of rational groups of genus s. If J/S is a direct sum of rational groups of genus $\mathbf{s}$, then J/S is separable, belongs to Γ_1 or to Γ_2 and all its elements $\neq 0$ have the genus $\mathbf{s}$. If finally J/S is separable, belongs to a class Γ_ν and all the elements $\neq 0$ of J/S have the genus $\mathbf{s}$, it follows from Corollary 4.4 that the conditions (a) to (c) of Theorem 8.4 are satisfied by J and its subgroup S, and J is therefore the direct sum of S and a completely reducible group.

Remark. Since there exist direct irreducible groups of rank 2 whose elements $\neq 0$ all have the same genus, the words "of genus $\mathbf{s}$" cannot be omitted in (b).

THEOREM 8.8. *If the subgroup S of the group J and the genus* $\mathbf{s}$ *satisfy*:

(a) $(S, \mathbf{s} < |x|)$ *is the intersection of S and* $(J, \mathbf{s} < |x|)$;

(b) $J(\mathbf{s})^* = (J, \mathbf{s} \leqq |x|)/(J, \mathbf{s} < |x|)$ *is a direct sum of rational groups of genus* $\mathbf{s}$;

then $(S, \mathbf{s} \leqq |x|)$ *is the direct sum of* $(S, \mathbf{s} < |x|)$ *and a completely reducible group* $S(\mathbf{s})$ *and* $r(S(\mathbf{s})) \leqq r(J(\mathbf{s})^*)$, *i.e.,* $S(\mathbf{s})$ *is isomorphic with a direct summand of* $J(\mathbf{s})^*$.

Proof.[16] By condition (b) and Corollary 8.7a, b there exist rational groups R_v of genus $\mathbf{s}$ such that

$$(J, \mathbf{s} \leqq |x|) = (J, \mathbf{s} < |x|) + \sum_v R_v .$$

It can be assumed that the indices v are the ordinal numbers, satisfying $0 < v < \sigma$. Put $J_v = (J, \mathbf{s} < |x|) + \sum_{\rho<v} R_\rho$ for $0 < v \leqq \sigma$ and S_v = intersection of $(S, \mathbf{s} \leqq |x|)$ and J_v. Then $J_1 = (J, \mathbf{s} < |x|)$, $S_1 = (S, \mathbf{s} < |x|)$ by (a), $J_{v+1} = J_v + R_v$, J_v = join of the groups J_ρ with $\rho < v$, if v is a limit ordinal, $J_\sigma = (J, \mathbf{s} \leqq |x|)$, $S_\sigma = (S, \mathbf{s} \leqq |x|)$.

Since S_{v+1}/S_v is isomorphic with a subgroup of R_v, it follows therefore by Corollary 8.2 that either $S_v = S_{v+1}$ or that S_{v+1}/S_v is a rational group of genus $\mathbf{s}$. Thus by Theorem 8.4 $S_{v+1} = S_v + T_v$, where either $T_v = 0$ or a rational group of genus $\mathbf{s}$.

[16] A similar proof for the complete reducibility of the subgroups of direct sums of infinite cyclic groups is due to L. Zippin.

The subgroup S' of $(S, \mathbf{s} \leqq | x |)$ which is generated by $(S, \mathbf{s} < | x |)$ and the groups T_v is the direct sum of these groups. $S_1 \leqq S'$. It can therefore be assumed that $S_\rho \leqq S'$ for every $\rho < v$. If $v = v' + 1$ is not a limit number, then $S_v = S_{v'} + T_{v'} \leqq S'$. If v is a limit ordinal, then S_v is the join of the groups S_ρ with $\rho < v$ and consequently $S_v \leqq S'$. Hence $S_\sigma \leqq S'$, i.e.,

$$(S, \mathbf{s} \leqq | x |) = (S, \mathbf{s} < | x |) + \sum_v T_v.$$

This completes the proof of the theorem.

Corollary 8.9. *Let J be a direct sum of rational groups of genus* $\mathbf{s}$. *Then the following three statements are equivalent.*

(a) *The subgroup S of J is a direct sum of rational groups of genus* $\mathbf{s}$.

(b) *All the elements $\neq 0$ of the subgroup S of J have in S the genus* $\mathbf{s}$.

(c) *The subgroup S of J is isomorphic with a direct summand of J.*

Proof. (a) implies (b). If (b) is satisfied, the conditions of Theorem 8.8 are satisfied, i.e., (b) implies (c). (c) implies (a), since every direct summand of J satisfies the conditions of Theorem 8.8.

Remark. This corollary implies in particular that every subgroup of a direct sum of infinite cyclic groups is a direct sum of infinite cyclic groups.

Corollary 8.10. *If the subgroup S of the group J and the genus* $\mathbf{s}$ *satisfy*

(a) $J(\mathbf{s})^* = (J, \mathbf{s} \leqq | x |)/(J, \mathbf{s} < | x |)$ *is a direct sum of rational groups of genus* $\mathbf{s}$;

(b) $(S, \mathbf{s} < | x |)$ *is the intersection of S and* $(J, \mathbf{s} < | x |)$;

(c) *the subgroup $S(\mathbf{s})^*$ of all those classes of $J(\mathbf{s})^*$ which contain elements of* $(S, \mathbf{s} \leqq | x |)$ *is a direct summand of* $J(\mathbf{s})^*$;

there exist subgroups U, V such that

$$(J, \mathbf{s} \leqq | x |) = U + V + (J, \mathbf{s} < | x |), \quad (S, \mathbf{s} \leqq | x |) = U + (S, \mathbf{s} < | x |)$$

and the elements of U represent exactly the classes of $S(\mathbf{s})^*$.

Proof. Since conditions (a), (b) and (c) imply the conditions of Theorem 8.8, it follows that

$$(S, \mathbf{s} \leqq | x |) = (S, \mathbf{s} < | x |) + U$$

where U is a direct sum of rational groups of genus $\mathbf{s}$. Furthermore, U represents exactly the classes of $S(\mathbf{s})^*$. By condition (c), $J(\mathbf{s})^* = T^* + S(\mathbf{s})^*$. By condition (a) and Corollary 8.7a, b it follows that

$$(J, \mathbf{s} \leqq | x |) = (J, \mathbf{s} < | x |) + J(\mathbf{s})$$

where $J(\mathbf{s})$ is a direct sum of rational groups of genus $\mathbf{s}$. If V is the subgroup of $J(\mathbf{s})$ which represents the classes of T^*, then

$$(J, \mathbf{s} \leqq | x |) = (J, \mathbf{s} < | x |) + U + V,$$

and this completes the proof.

Note that the above groups U and V are direct sums of rational groups of genus $\mathbf{s}$. Furthermore, it is a consequence of Corollary 4.4 that condition (c)

is satisfied, if $S(\mathbf{s})^*$ is a closed subgroup of finite rank of $J(\mathbf{s})^*$ or if $J(\mathbf{s})^*$ is complete and $S(\mathbf{s})^*$ a closed subgroup.

9. **Partial reducibility.**

NOTATION. *If J is any group, then*
$\mathbf{A}(J)$ *is the set of all the genera* $\mathbf{s}$ *such that*

$$J(\mathbf{s})^{**} = (J, |x| \not< \mathbf{s})/(J, |x| \not\leqq \mathbf{s}) \neq 0;$$

$\mathbf{B}(J)$ *is the set of all the genera* $\mathbf{s}$ *such that*

$$J(\mathbf{s})^* = (J, \mathbf{s} \leqq |x|)/(J, \mathbf{s} < |x|) \neq 0;$$

and $\mathbf{C}(J)$ *is the set of all the genera* $\mathbf{s}$ *such that elements of genus* $\mathbf{s}$ *exist in* J.

$\mathbf{A}(J) \leqq \mathbf{B}(J)$, since there exists by (2.4b) a homomorphism of $J(\mathbf{s})^*$ upon the whole group $J(\mathbf{s})^{**}$, and it follows from (2.4b) that $\mathbf{A}(J) = \mathbf{B}(J)$, if $(J, \mathbf{s} < |x|)$ is always the intersection of $(J, \mathbf{s} \leqq |x|)$ and $(J, |x| \not\leqq \mathbf{s})$.

$\mathbf{B}(J) \leqq \mathbf{C}(J)$, since elements of $(J, \mathbf{s} \leqq |x|)$ which are not contained in $(J, \mathbf{s} < |x|)$ have in J the genus $\mathbf{s}$.

CHAIN CONDITION. *If* $\mathbf{s}_1 \leqq \mathbf{s}_2 \leqq \cdots \leqq \mathbf{s}_i \leqq \cdots$ *is an ascending chain of genera* $\mathbf{s}_i$ *in the set* $\mathbf{G}$ *of genera, almost all the* $\mathbf{s}_i$ *are equal.*

If the set $\mathbf{G}$ of genera satisfies the chain condition, every subset of $\mathbf{G}$ satisfies the chain condition and every subset of $\mathbf{G}$ contains a "greatest" element $\mathbf{g}$ such that $\mathbf{g} \not< \mathbf{s}$ for every genus $\mathbf{s}$ in $\mathbf{G}$ and the chain condition is conversely a consequence of this condition.

(9.1) *If $r(J)$ is finite, the chain condition is satisfied in* $\mathbf{C}(J)$ (*and therefore in* $\mathbf{A}(J)$ *and in* $\mathbf{B}(J)$).

For if $|b| = \mathbf{s} < \mathbf{t}$, then $(J, \mathbf{t} \leqq |x|) < (J, \mathbf{s} \leqq |x|)$ and since $(J, \mathbf{g} \leqq |x|)$ is a closed subgroup of J and $r(J)$ is finite, this implies $r((J, \mathbf{t} \leqq |x|)) < r((J, \mathbf{s} \leqq |x|))$, and thus the chain condition holds in $\mathbf{C}(J)$, if $r(J)$ is finite.[17]

It may happen that $r(J)$ is finite and $\mathbf{B}(J)$ is infinite. This is shown by the following

Example 9.2. Let $p_1, p_2, \cdots$ be an enumeration of the prime numbers and J a group of rank 2 which contains two independent elements b' and b'' such that the p-value of $m(b' + p_i b'')$ is 0, if $p \neq p_i$, and infinite, if $p = p_i$. If $\mathbf{p}_i$ is the genus of p_i^∞, then $0 = (J, \mathbf{p}_i < |x|)$, and $(J, \mathbf{p}_i \leqq |x|)$ consists of all the rational multiples of $b' + p_i b''$, i.e., $\mathbf{B}(J)$ contains every $\mathbf{p}_i$ and is therefore infinite.

THEOREM 9.3. *If the chain condition is satisfied in* $\mathbf{B}(J)$, *then the following three propositions are equivalent.*

(a) *J is partially reducible.*

(b) *J has the following properties:*

[17] Another consequence of this argument is that also the "descending chain condition" holds in $\mathbf{C}(J)$.

(b1) *for every genus* $\mathbf{s}$, $(J, \mathbf{s} < |x|)$ *is a direct summand of* $(J, \mathbf{s} \leqq |x|)$;

(b2) *for every genus* $\mathbf{s}$, $(J, \mathbf{s} < |x|)$ *is the intersection of* $(J, \mathbf{s} \leqq |x|)$ *and* $(J, |x| \nleqq \mathbf{s})$;

(b3) *corresponding to every element* $b \neq 0$ *in* J *there exists at least one genus* $\mathbf{s}$ *such that* $b \equiv 0 \bmod (J, |x| \nless \mathbf{s})$, $b \not\equiv 0 \bmod (J, |x| \nleqq \mathbf{s})$;

(b4) *to every element* b *of* J *there exists at most a finite number of genera* $\mathbf{s}$ *such that* $b \equiv 0 \bmod (J, |x| \nless \mathbf{s})$, $b \not\equiv 0 \bmod (J, |x| \nleqq \mathbf{s})$.

(c) *If for every genus* $\mathbf{s}$ *the subgroup* $J(\mathbf{s})$ *of* J *satisfies*

$$(\mathrm{cs}) \qquad (J, \mathbf{s} \leqq |x|) = (J, \mathbf{s} < |x|) + J(\mathbf{s}),$$

then J *is the direct sum of these groups* $J(\mathbf{s})$.[18]

Proof. 1. If J is partially reducible, it follows from (2.6) and (2.7) that J fulfills the conditions (b1) to (b4).

2. If J satisfies (b2) and $J(\mathbf{s})$ is a solution of (cs), then $J(\mathbf{s})$ is by (2.4) also a solution of

$$(\bar{\mathrm{c}}\mathrm{s}) \qquad (J, |x| \nless \mathbf{s}) = J(\mathbf{s}) + (J, |x| \nleqq \mathbf{s}).$$

3. Suppose that J satisfies (b2) and that, for every genus $\mathbf{s}$, $J(\mathbf{s})$ is a solution of (cs). If $b_1, \cdots, b_k$ are a finite number of elements $\neq 0$ which belong to different groups $J(\mathbf{s})$, there exists one among them, say b_1, such that $|b_i| \nless |b_1|$. Then

$$b = \sum_{i=1}^{k} b_i \equiv 0 \bmod (J, |x| \nless |b_1|), \qquad b \equiv b_1 \not\equiv 0 \bmod (J, |x| \nleqq |b_1|),$$

since, as proved in 2, $J(|b_1|)$ is a solution of $(\bar{\mathrm{c}} \mid b_1 |)$. Hence $b \neq 0$ and this implies

(9.3.3) If J satisfies (b2) and if, for every genus $\mathbf{s}$, $J(\mathbf{s})$ is a solution of (cs), then the subgroup of J, generated by the groups $J(\mathbf{s})$, is their direct sum.

4. Suppose that J satisfies the conditions (b1) to (b4) and that, for every genus $\mathbf{s}$, $J(\mathbf{s})$ is a solution of (cs).

If $b \neq 0$ is an element of J, denote by $\mathbf{A}(b)$ the set of all the genera $\mathbf{s}$ such that $b \equiv 0 \bmod (J, |x| \nless \mathbf{s})$, $b \not\equiv 0 \bmod (J, |x| \nleqq \mathbf{s})$. Denote by J' the direct sum of the groups $J(\mathbf{s})$ (which exists by (9.3.3) in J).

Since every $J(\mathbf{s})$ is also a solution of $(\bar{\mathrm{c}}\mathrm{s})$, there exists to every element $b \neq 0$ and to every genus $\mathbf{s}$ in $\mathbf{A}(b)$ a uniquely determined element $b(\mathbf{s})$ in $J(\mathbf{s})$ such that $b \equiv b(\mathbf{s}) \bmod (J, |x| \nleqq \mathbf{s})$. Since by (b4) the set $\mathbf{A}(b)$ is finite, it is possible to form the sum b' of all the elements $b(\mathbf{s})$ with $\mathbf{s}$ in $\mathbf{A}(b)$.

Put $b'' = b - b'$ and let $\mathbf{w}$ be a genus such that $\mathbf{s} \nless \mathbf{w}$ for every $\mathbf{s}$ in $\mathbf{A}(b)$. Since every $b(\mathbf{s}) \neq 0$ and therefore every $b(\mathbf{s})$ is an element of genus $\mathbf{s}$, this implies that $b \equiv b'' \bmod (J, |x| \nless \mathbf{w})$. If $\mathbf{w} \neq \mathbf{s}$ for every $\mathbf{s}$ in $\mathbf{A}(b)$, then $b(\mathbf{s}) \equiv 0 \bmod (J, |x| \nleqq \mathbf{w})$, i.e., $b'' \equiv b \equiv 0 \bmod (J, |x| \nleqq \mathbf{w})$ and $\mathbf{w}$ does not belong to $\mathbf{A}(b'')$. If $\mathbf{w}$ is an element of $\mathbf{A}(b)$, then $b \equiv b(\mathbf{w}) \bmod (J, |x| \nleqq \mathbf{w})$

[18] And there exist solutions of the equations (cs).

and therefore $b'' \equiv 0 \bmod (J, |x| \nleqq \mathbf{w})$, i.e., $\mathbf{w}$ does not belong to $\mathbf{A}(b'')$. Thus it has been proved that $b \equiv b'' \bmod J'$, and every element of $\mathbf{A}(b'')$ is a true multiple of at least one genus in $\mathbf{A}(b)$.

5. Suppose finally that the chain condition is satisfied by $\mathbf{B}(J)$ and that J satisfies (b1) to (b4). There exist solutions $J(\mathbf{s})$ of (cs), their direct sum J' is a subgroup of J and every element b of J is congruent mod J' to an element b' of J such that every genus in $\mathbf{A}(b')$ is a true multiple of at least one genus in $\mathbf{A}(b)$. Since $\mathbf{A}(u)$ is vacuous if, and only if, $u = 0$, and since every $\mathbf{A}(u)$ is finite and a subset of $\mathbf{B}(J) = \mathbf{A}(J)$, it follows from the chain condition that $J = J'$, i.e., (b) implies (c).

6. Since the decomposition of J into the solutions $J(\mathbf{s})$ of (cs) is a partial reduction of J, (c) implies (a).

That (a) is not a consequence of (b), if the chain condition in $\mathbf{B}(J)$ is not satisfied, is shown by the following

Example 9.4. Let $p_1, p_2, \cdots$ be an enumeration of the prime numbers, R_i a rational group of genus $\mathbf{r}_i = \left| \prod_{j=1}^{i} p_j^{\infty} \right|$, J the additive group of all the vectors whose i-th coördinate is an element of R_i.

The subgroup of all the vectors of J whose coördinates except the i-th are 0 is isomorphic with R_i and may be identified with R_i.

Since $\mathbf{r}_{i-1} < \mathbf{r}_i$, a vector v of J whose i-th coördinate is the first coördinate $\neq 0$ has the genus $\mathbf{r}_i$ in J. Thus $\mathbf{C}(J)$ consists exactly of the genera $\mathbf{r}_i$,

$$(J, \mathbf{r}_i \leqq |x|) = (J, |x| \nless r_i) = R_i + (J, \mathbf{r}_i < |x|) = (J, \mathbf{r}_{i-1} < |x|),$$

and therefore $\mathbf{A}(J) = \mathbf{B}(J) = \mathbf{C}(J)$.

J satisfies all the conditions (b1) to (b4). But since J contains a continuum of elements and the direct sum of the groups R_i is countable, J is neither completely nor partially reducible.[19]

That (c) is not a consequence of (a)—and even not of complete reducibility—if the chain condition in $\mathbf{B}(J)$ does not hold is shown by

Example 9.5. Denote by $\mathbf{s}_1 < \mathbf{s}_2 < \cdots < \mathbf{s}_i < \cdots$ some ascending chain of genera, by J_i a rational group of genus $\mathbf{s}_i$ and by J the direct sum of the groups J_i. If $b_i \neq 0$ is an element of J_i, these elements b_i form a basis of J. Put $w_i = b_i + b_{i+1}$ and let $\overline{w}_i$ be the closed subgroup of J, generated by w_i. The groups $\overline{w}_i$ form a complete set of solutions $\neq 0$ of the equations (cs), if $m(b_i) \mid m(b_{i+1})$, but their direct sum W does not contain b_1 and is $< J$.

That (b3) is not a consequence of the chain condition in $\mathbf{B}(J)$ and the conditions (b1), (b2), (b4), is shown by

Example 9.6. Let $p_1, p_2, \cdots$ be an enumeration of the prime numbers, J' the group generated by the elements $(i_1, \cdots, i_k)$ for $k = 1, 2, \cdots, i_j = 0, 1$, satisfying the relations $(i_1, \cdots, i_k) = (i_1, \cdots, i_k, 0) + (i_1, \cdots, i_k, 1)$ for

[19] The separability of this group J will be proved in section 11.

$0 < k$. J is the smallest group between J' and some complete group such that the p-value of $m((i_1, \cdots, i_k) < J)$ is $\begin{cases} 0, \text{ if } p \neq p_{2^j+i_j} \text{ for every } j \leqq k \\ \infty, \text{ if } p = p_{2^j+i_j} \text{ for some } j \leqq k. \end{cases}$

Every element $(i_1, \cdots, i_k)$ is therefore in J the sum of two elements whose genera are true multiples of $|(i_1, \cdots, i_k)|$. Consequently $(J, \mathbf{s} < |x|) = (J, \mathbf{s} \leqq |x|)$ and $(J, |x| \nleqq \mathbf{s}) = (J, |x| \nless \mathbf{s})$ for every genus $\mathbf{s}$, i.e., $\mathbf{A}(J)$ and $\mathbf{B}(J)$ are both vacuous. Thus J is a countable group, the chain condition satisfied in $\mathbf{A}(J)$ and in $\mathbf{B}(J)$, conditions (b1), (b2) and (b4) are satisfied in J. But (b3) is not satisfied.

COROLLARY 9.7. *If $r(J)$ is finite, the following propositions are equivalent.*

(a) *J is partially reducible.*

(b) *For every genus $\mathbf{s}$, $(J, \mathbf{s} < |x|)$ is a direct summand of $(J, \mathbf{s} \leqq |x|)$ and*

$$r(J(\mathbf{s})^*) = r(J(\mathbf{s})^{**}/F(J(\mathbf{s})^{**})).$$

(c) *For every genus $\mathbf{s}$, $(J, \mathbf{s} < |x|)$ is a direct summand of $(J, \mathbf{s} \leqq |x|)$ and is the intersection of $(J, \mathbf{s} \leqq |x|)$ and $(J, |x| \nless \mathbf{s})$.*

(d) *If for every genus $\mathbf{s}$ the subgroup $J(\mathbf{s})$ of J satisfies*

$$\text{(ds)} \qquad (J, \mathbf{s} \leqq |x|) = J(\mathbf{s}) \dotplus (J, \mathbf{s} < |x|),$$

then J is the direct sum of these groups $J(\mathbf{s})$.[20]

Proof. 1. If J is partially reducible, then $J(\mathbf{s})^*$ and $J(\mathbf{s})^{**}$ are by (2.6) isomorphic groups and thus (b) is a consequence of (a) and (2.6).

2. Suppose that $r(J)$ is finite and that J satisfies (b). There exists by (2.4) a homomorphism of $J(\mathbf{s})^*$ upon $J(\mathbf{s})^{**}$ and therefore there exists a homomorphism α of $J(\mathbf{s})^*$ upon $J(\mathbf{s})^{**}/F(J(\mathbf{s})^{**})$. If W^* is the subgroup of $J(\mathbf{s})^*$ which is mapped upon 0 by α, then $J(\mathbf{s})^*/W^*$ and $J(\mathbf{s})^{**}/F(J(\mathbf{s})^{**})$ are isomorphic. Since therefore W^* is a closed subgroup of $J(\mathbf{s})^*$ and

$$r(J(\mathbf{s})^*) = r(W^*) + r(J(\mathbf{s})^{**}/F(J(\mathbf{s})^{**})) = r(W^*) + r(J(\mathbf{s})^*),$$

and since all the ranks are finite, it follows that $W^* = 0$. The homomorphism of $J(\mathbf{s})^*$ upon $J(\mathbf{s})^{**}$ is therefore an isomorphism and it follows from (2.4) that J satisfies (c).

3. Suppose that $r(J)$ is finite and that J satisfies (c). Then every subgroup $(J, \mathbf{s} \leqq |x|)$ has finite rank and satisfies (c).

If $r(J) = 1$, then (c) and (d) are satisfied. It can therefore be assumed that the group J' satisfies (d), if $r(J') < r(J)$ and if J' satisfies (c).

Suppose now that the subgroups $J(\mathbf{s})$ of J are solutions of (ds) and that J' is the subgroup of J generated by the subgroups $J(\mathbf{s})$. Then it follows from (c) and (9.3.3) that J' is the direct sum of the groups $J(\mathbf{s})$.

Let now $b \neq 0$ be an element of J and $\mathbf{g}$ its genus in J. Then $(J, \mathbf{t} \leqq |x|) < J$ and therefore $r((J, \mathbf{t} \leqq |x|)) < r(J)$ for every genus $\mathbf{g} < \mathbf{t}$. Consequently

[20] And there exist solutions of the equations (ds).

every $(J, \mathbf{t} \leqq |x|)$ with $\mathbf{g} < \mathbf{t}$ is the direct sum of the groups $J(\mathbf{s})$ with $\mathbf{t} \leqq \mathbf{s}$. Since $(J, \mathbf{g} < |x|)$ is the join of the groups $(J, \mathbf{t} \leqq |x|)$ with $\mathbf{g} < \mathbf{t}$, $(J, \mathbf{g} < |x|)$ is the direct sum of the groups $J(\mathbf{s})$ with $\mathbf{g} < \mathbf{s}$ and $(J, \mathbf{g} \leqq |x|)$ is therefore the direct sum of the groups $J(\mathbf{s})$ with $\mathbf{g} \leqq \mathbf{s}$, i.e., b is contained in J', i.e., $J = J'$ is the direct sum of the groups $J(\mathbf{s})$.

4. (d) implies (a), since the decomposition of J into the solutions of (ds) is a partial reduction of J.

Corollary 9.8. *Suppose that* $\mathbf{B}(J)$ *satisfies*

(0) *If* $\mathbf{s}$ *and* $\mathbf{t}$ *are two genera in* $\mathbf{B}(J)$, *then either* $\mathbf{s} < \mathbf{t}$ *or* $\mathbf{s} = \mathbf{t}$ *or* $\mathbf{t} < \mathbf{s}$.

(I) *If the chain condition is satisfied in* $\mathbf{B}(J)$, *the propositions* (a) to (c) *of Theorem* 9.3 *are true in* J *if, and only if,*

(I′) *for every genus* $\mathbf{s}$, $(J, \mathbf{s} < |x|)$ *is a direct summand of* $(J, \mathbf{s} \leqq |x|)$,

(I″) *elements* $b \neq 0$ *are not contained in* $(J, |b| < |x|)$.

(II) *If* $r(J)$ *is finite, the propositions* (a) *to* (d) *of Corollary* 9.7 *are true in* J *if, and only if, for every genus* $\mathbf{s}$, $(J, \mathbf{s} < |x|)$ *is a direct summand of* $(J, \mathbf{s} \leqq |x|)$.

Proof. If J is partially reducible and satisfies (0), $b = \sum_{i=1}^{k} b_i$ is an element $\neq 0$ of J and the elements $b_i \neq 0$ belong to different components of a smallest partial reduction of J, then $|b_1| < \cdots < |b_k|$ and $|b| = |b_1|$, i.e., (I′) and (I″) are satisfied by J.

If conversely (0), (I′) and (I″) are satisfied in J, then $\mathbf{A}(J) = \mathbf{B}(J) = \mathbf{C}(J)$, $(J, \mathbf{s} \leqq |x|) = (J, |x| \nless s)$, $(J, \mathbf{s} < |x|) = (J, |x| \nleqq \mathbf{s})$ and conditions (b1) to (b4) of Theorem 9.3 are therefore satisfied by J.

If $r(J)$ is finite and there exists for every genus $\mathbf{s}$ a solution $J(\mathbf{s})$ of (ds) in Corollary 9.7, the subgroup J' of J, generated by these groups $J(\mathbf{s})$, is by (0) their direct sum. If $b \neq 0$ is an element of genus $\mathbf{g}$, then $(J, \mathbf{g} < |x|)$ is the join of the groups $(J, \mathbf{t} \leqq |x|)$ with $\mathbf{g} < \mathbf{t}$. Then $r((J, \mathbf{t} \leqq |x|)) < r(J)$ and it can therefore be assumed that $(J, \mathbf{t} \leqq |x|)$ is the direct sum of the groups $J(\mathbf{s})$ with $\mathbf{g} < \mathbf{t} \leqq \mathbf{s}$, i.e., $(J, \mathbf{g} < |x|)$ and consequently $(J, \mathbf{g} \leqq |x|)$ are subgroups of J', i.e., b is an element of J', i.e., $J = J'$ is the direct sum of the groups $J(\mathbf{s})$.

Note that the examples 9.4 to 9.6 satisfy (0).

Corollary 9.9. *Suppose that the chain condition is satisfied in* $\mathbf{B}(J)$. *Then* J *is partially reducible if, and only if,*

(a) *for every genus* $\mathbf{s}$, $(J, \mathbf{s} < |x|)$ *is a direct summand of* $(J, \mathbf{s} \leqq |x|)$;

(b) *for every genus* $\mathbf{s}$, $(J, \mathbf{s} < |x|)$ *is the join of the groups* $(J, \mathbf{t} \leqq |x|)$ *with* $\mathbf{s} < \mathbf{t}$ *and* $\mathbf{t}$ *in* $\mathbf{B}(J)$;

(c) *for every genus* $\mathbf{s}$, $(J, \mathbf{s} < |x|)$ *is the intersection of*$(J, \mathbf{s} \leqq |x|)$ *and* $(J, |x| \nleqq \mathbf{s})$.

Proof. The necessity of the conditions (a) and (b) is a consequence of (2.6) and the necessity of (c) a consequence of (2.7).

If conversely (a) to (c) and the chain condition in $\mathbf{B}(J)$ are satisfied, then let, for every genus $\mathbf{s}$, $J(\mathbf{s})$ be a subgroup of J such that $(J, \mathbf{s} \leqq |x|) = J(\mathbf{s}) + (J, \mathbf{s} < |x|)$. The joingroup J' of these groups is by (c) and (9.3.3) their direct sum.

Let **W** be the set of all those genera $\mathbf{s}$ in $\mathbf{B}(J)$ such that $(J, \mathbf{s} \leqq |x|) \nleqq J'$. If **W** is not vacuous, there exists from the chain condition a "greatest" genus $\mathbf{w}$ in **W**. $(J, \mathbf{w} < |x|)$ is by (b) the join of the groups $(J, \mathbf{t} \leqq |x|)$ with $\mathbf{w} < \mathbf{t}$ and $\mathbf{t}$ in $B(J)$ and since these genera $\mathbf{t}$ are not contained in **W**, it follows that $(J, \mathbf{w} < |x|) \leqq J'$ and consequently that $(J, \mathbf{w} \leqq |x|) \leqq J'$. Thus **W** is vacuous, i.e., $(J, \mathbf{s} \leqq |x|) \leqq J'$, if $\mathbf{s}$ in $\mathbf{B}(J)$. If $\mathbf{z}$ is the genus of the number 1, then $J = (J, \mathbf{z} \leqq |x|) = J(\mathbf{z}) + (J, \mathbf{z} < |x|)$ and by (b) it follows that $(J, \mathbf{z} < |x|) \leqq J'$, i.e., that $J = J'$ is the direct sum of the groups $J(\mathbf{s})$.

THEOREM 9.10. *If the subgroup S of the partially reducible group J satisfies*

(a) *the chain condition in* $\mathbf{B}(S)$;

(b) *for every genus* $\mathbf{s}$, $(S, \mathbf{s} < |x|)$ *is a direct summand of* $(S, \mathbf{s} \leqq |x|)$;

(c) $(S, f(x))$ *is the intersection of S and $(J, f(x))$ for every* $f(x) = (\mathbf{s} \leqq |x|)$, $(\mathbf{s} < |x|)$, $(|x| \nless \mathbf{s})$, $(|x| \nleqq \mathbf{s})$;

then S is partially reducible and every $S(\mathbf{s})^*$ is isomorphic with a subgroup of $J(\mathbf{s})^*$.

Proof. J satisfies the conditions (a)–(c) of Corollary 9.9 since J is partially reducible. Thus conditions (a)–(c) of Theorem 9.10 imply that S satisfies the conditions of Corollary 9.9 and S is therefore partially reducible.

COROLLARY 9.11. *Suppose that S is a subgroup of the partially reducible group J and that the chain condition is satisfied in* $\mathbf{B}(J)$. *Then S is a direct summand of J if, and only if,*

(a) *for every genus* $\mathbf{s}$, $(S, \mathbf{s} < |x|)$ *is the intersection of S and* $(J, \mathbf{s} < |x|)$;

(b) *for every genus* $\mathbf{s}$, *the elements of* $(S, \mathbf{s} \leqq |x|)$ *represent a direct summand of* $J(\mathbf{s})^*$;

(c) *S is partially reducible.*

Proof. Suppose first that S is a direct summand of J. Then (a) and (b) are satisfied, since $(S, f(x))$ is a direct summand of $(J, f(x))$ (for the discussed functions $f(x)$), and since therefore the conditions (a)–(c) of Theorem 9.10 are satisfied in S and J, S is also partially reducible.

Suppose now that the subgroup S of J satisfies the conditions (a)–(c). Then there exists a smallest partial reduction of S, $S = \sum_{\mathbf{s}} S(\mathbf{s})$, and the subgroup $S(\mathbf{s})$ of S represents exactly the subgroup $S(\mathbf{s})^*$ of $J(\mathbf{s})^*$ which is represented by elements of $(S, \mathbf{s} \leqq |x|)$. By (b), $J(\mathbf{s})^* = S(\mathbf{s})^* + T(\mathbf{s})^*$. There exists furthermore a partial reduction of J and if $J(\mathbf{s})$ satisfies $(J, \mathbf{s} \leqq |x|) = J(\mathbf{s}) + (J, \mathbf{s} < |x|)$, denote by $T(\mathbf{s})$ the subgroup of $J(\mathbf{s})$ which represents the classes of $T(\mathbf{s})^*$. Then

$$(J, \mathbf{s} \leqq |x|) = S(\mathbf{s}) + T(\mathbf{s}) + (J, \mathbf{s} < |x|),$$

and therefore by Theorem 9.3

$$J = \sum_{\mathbf{s}} (S(\mathbf{s}) + T(\mathbf{s})) = \sum_{\mathbf{s}} S(\mathbf{s}) + \sum_{\mathbf{s}} T(\mathbf{s}) = S' + T,$$

and since the chain condition is also satisfied in $\mathbf{B}(S) \leqq \mathbf{B}(J)$, it follows from Theorem 9.3 that $S = S'$ is a direct summand of J.

Note that it is possible to substitute for condition (c) of the Corollary 9.11 the condition (c) of Theorem 9.10.

10. **Complete reducibility.** It is the object of this section to combine the results of the two preceding sections.

THEOREM 10.1. *The group J is completely reducible if, and only if,*

(a) *for every genus* $\mathbf{s}$, $J(\mathbf{s})^*$ *is a direct sum of rational groups of genus* $\mathbf{s}$;

(b) *for every genus* $\mathbf{s}$, $(J, \mathbf{s} < |x|)$ *is the intersection of* $(J, \mathbf{s} \leqq |x|)$ *and* $(J, |x| \nleqq \mathbf{s})$;

(c) *the set* $\mathbf{A}(b)$ *of all the genera* $\mathbf{s}$ *such that*

$$b \equiv 0 \bmod (J, |x| \nless \mathbf{s}), \qquad b \not\equiv 0 \bmod (J, |x| \nleqq \mathbf{s})$$

is for every element $b \neq 0$ of J finite and not vacuous;

(d) *J is the direct sum of groups J_ν such that the chain condition is satisfied in every* $\mathbf{B}(J_\nu)$.

This is a consequence of Theorem 9.3 and Corollary 8.7, since the conditions (a)–(c) are satisfied in a direct sum if, and only if, they are satisfied in all the direct summands.

COROLLARY 10.2. *The group J of finite rank is completely reducible if, and only if, for every genus* $\mathbf{s}$

(a) $J(\mathbf{s})^*$ *is a direct sum of rational groups of genus* $\mathbf{s}$;

(b) $r(J(\mathbf{s})^*) = r(J(\mathbf{s})^{**}/F(J(\mathbf{s})^{**}))$.

This is a consequence of Corollaries 9.7 and 8.7.

By Corollary 9.7 it is possible to substitute for (b) the condition

(b′) $(J, \mathbf{s} < |x|)$ *is the intersection of* $(J, \mathbf{s} \leqq |x|)$ *and* $(J, |x| \nleqq \mathbf{s})$.

THEOREM 10.3. *If the subgroup S of the completely reducible group J satisfies*

(a) *the chain condition holds in* $\mathbf{B}(S)$;

(b) $(S, f(x))$ *is the intersection of S and $(J, f(x))$ for every* $f(x) = (\mathbf{s} \leqq |x|)$, $(\mathbf{s} < |x|)$, $(|x| \nless \mathbf{s})$, $(|x| \nleqq \mathbf{s})$;

then S is completely reducible and isomorphic with a direct summand of J.

This is a consequence of Theorems 9.10 and 8.8.

An obvious consequence of Theorem 10.3 is the

COROLLARY 10.4. *Every direct summand D of the completely reducible group J such that the chain condition holds in $\mathbf{B}(D)$ is completely reducible.*

COROLLARY 10.5. *Suppose that S is a subgroup of the completely reducible group J and that the chain condition holds in $\mathbf{B}(J)$. Then S is a direct summand of J if, and only if,*

(a) *for every genus* $\mathbf{s}$, $(S, \mathbf{s} < |x|)$ *is the intersection of S and* $(J, \mathbf{s} < |x|)$;

(b) *for every genus* $\mathbf{s}$, $(S, |x| \nleqq \mathbf{s})$ *is the intersection of S and* $(J, |x| \nleqq \mathbf{s})$;

(c) *for every genus $\mathbf{s}$ the elements of $(S, \mathbf{s} \leqq |x|)$ represent a direct summand of* $J(\mathbf{s})^*$.

This is a consequence of Corollary 8.10, Theorem 9.3 and Corollary 9.9 (since every condition used is satisfied in J).

COROLLARY 10.6. *Let J be a completely reducible group. Then every subgroup S of J, satisfying for every genus* $\mathbf{s}$ *the three conditions*

(a) *$(S, \mathbf{s} < |x|)$ is the intersection of S and $(J, \mathbf{s} < |x|)$;*

(b) *$(S, |x| \nleqq \mathbf{s})$ is the intersection of S and $(J, |x| \nleqq \mathbf{s})$;*

(c) *the elements of $(S, \mathbf{s} \leqq |x|)$ represent a closed subgroup of $J(\mathbf{s})^*$;*

is a direct summand of J if, and only if,

(1) *the chain condition is satisfied in $\mathbf{B}(J)$;*

(2) *for every genus $\mathbf{s}$, $J(\mathbf{s})^*$ is complete or $r(J(\mathbf{s})^*)$ is finite.*

Proof. The necessity of (1) may be derived from Example 9.5 and the necessity of (2) is a consequence of Corollary 3.6. If (1) and (2) are satisfied, every subgroup S, satisfying (a)–(c), by Corollary 3.6 satisfies the conditions of Corollary 10.5 and is therefore a direct summand of J.

11. Separability.

LEMMA 11.1. *The group J is separable if, and only if,*

(a) *every finite subset of J is contained in a partially reducible direct summand of J;*

(b) *$J(\mathbf{s})^* = (J, \mathbf{s} \leqq |x|)/(J, \mathbf{s} < |x|)$ is separable for every genus* $\mathbf{s}$.

Proof. Suppose first that J is separable. Then (a) is satisfied, since every complete reduction of a group is also a partial reduction. If F^* is a finite subset of $J(\mathbf{s})^*$, F a subset of J representing F^*, then F is contained in a completely reducible direct summand D of J. In particular, $(D, \mathbf{s} \leqq |x|)$ is the direct sum of $(D, \mathbf{s} < |x|)$ and rational groups of genus $\mathbf{s}$, $(D, \mathbf{s} \leqq |x|)$ is a direct summand of $(J, \mathbf{s} \leqq |x|)$ and $(D, \mathbf{s} < |x|)$ a direct summand of $(J, \mathbf{s} < |x|)$. The elements of $(D, \mathbf{s} \leqq |x|)$ represent therefore a direct summand of $J(\mathbf{s})^*$ which contains F^* and is completely reducible, i.e., $J(\mathbf{s})^*$ is separable.

Assume now that the conditions (a) and (b) are satisfied by J. Then it follows from (2.7) that every element $\neq 0$ of $J(\mathbf{s})^*$ has the genus $\mathbf{s}$ in $J(\mathbf{s})^*$ and consequently every closed subgroup of finite rank of $J(\mathbf{s})^*$ is by Corollary 4.4 a direct summand of $J(\mathbf{s})^*$ and by Corollary 4.3 a direct sum of a finite number of rational groups of genus $\mathbf{s}$. If now F is a finite subset of J, then F is contained in a partially reducible direct summand D of J. If $D = \sum_{\mathbf{s}} D(\mathbf{s})$ is a smallest partial reduction of D, the components of elements of F in $D(\mathbf{s})$ form a finite subset $F(\mathbf{s})$ of $D(\mathbf{s})$. Since $D(\mathbf{s})$ represents exactly a direct summand of $J(\mathbf{s})^*$, it follows from the above that every closed subgroup of finite rank of $D(\mathbf{s})$ is a completely reducible direct summand of $D(\mathbf{s})$. The closed subgroup $\overline{F(\mathbf{s})}$ of $D(\mathbf{s})$, generated by $F(\mathbf{s})$, is therefore a completely reducible direct summand of $D(\mathbf{s})$ and F is contained in a completely reducible direct summand of J, i.e., J is separable.

A consequence of this proof is the

COROLLARY 11.2. *The group J is separable if, and only if,*

(a) *every finite subset of J is contained in a partially reducible direct summand of J;*

(b′) *every closed subgroup of finite rank of* $J(\mathbf{s})^*$ *is a direct summand of* $J(\mathbf{s})^*$ (*and a direct sum of rational groups of genus* $\mathbf{s}$).

THEOREM 11.3. *Suppose that the chain condition is satisfied in* $\mathbf{B}(J)$. *Then the group* J *is separable if, and only if,*

(a) *for every genus* $\mathbf{s}$, $(J, \mathbf{s} < |x|)$ *is the intersection of* $(J, \mathbf{s} \leqq |x|)$ *and* $(J, |x| \nleqq \mathbf{s})$;

(b) *for every genus* $\mathbf{s}$, $J(\mathbf{s})^*$ *is a separable group and all the elements* $\neq 0$ *of* $J(\mathbf{s})^*$ *have the genus* $\mathbf{s}$ *in* $J(\mathbf{s})^*$;

(c) *for every genus* $\mathbf{s}$, *every subgroup* S *satisfying*

(c1) *every element* $\neq 0$ *of* S *has the genus* $\mathbf{s}$ *in* J,

(c2) $r(S)$ *is finite,*

(c3) S *is a direct summand of* $(J, |x| \nless \mathbf{s})$,

is a direct summand of J;

(d) *the set* $\mathbf{A}(b)$ *of the genera* $\mathbf{s}$ *such that*

$$b \equiv 0 \bmod (J, |x| \nless \mathbf{s}), \qquad b \not\equiv 0 \bmod (J, |x| \nleqq \mathbf{s})$$

is for every element $b \neq 0$ *in* J *finite and not vacuous.*

Proof. 1. If J is separable, then J satisfies (a), (b) and (d) by Lemma 11.1 and (2.7). If furthermore the subgroup S of J satisfies (c1)–(c3), then S is by Theorem 4.2 a direct summand of J.

2. If the chain condition is satisfied in $\mathbf{B}(J)$ and J fulfills (a)–(d), then the same holds true for every direct summand of J. Thus it is sufficient to prove: if J satisfies (a)–(d) and if the chain condition holds in $\mathbf{B}(J)$, every element of J is contained in a completely reducible direct summand.

3. Suppose that J satisfies the conditions (a)–(c) and that F is a primitive set in J (see Definition 5.1).

Assume first that F consists of exactly one element b and that $\mathbf{s}$ is the genus of b. Denote by $\bar{b}$ the closed subgroup of J, generated by b, and by $\bar{b}^*$ the subgroup of $J(\mathbf{s})^*$, represented by elements of $\bar{b}$. Since b is a primitive element of its genus $\mathbf{s}$, $\bar{b}^*$ is a closed subgroup of rank one of $J(\mathbf{s})^*$ and by (b) and Corollary 4.4 therefore a direct summand of $J(\mathbf{s})^*$, i.e., $J(\mathbf{s})^* = \bar{b}^* + T^*$. If T is the subgroup of $(J, \mathbf{s} \leqq |x|)$ which contains $(J, \mathbf{s} < |x|)$ and satisfies $T^* = T/(J, \mathbf{s} < |x|)$, then $(J, \mathbf{s} \leqq |x|) = \bar{b} + T$, since the elements of $\bar{b}$ represent exactly the classes of $\bar{b}^*$. If T' is the subgroup of $(J, |x| \nless \mathbf{s})$, generated by T and $(J, |x| \nleqq \mathbf{s})$, it follows from (a) and (2.4) that $(J, |x| \nless \mathbf{s}) = \bar{b} + T'$, and (c) implies therefore that $\bar{b}$ is a direct summand of J.

Suppose now that F contains k elements $b_1, \cdots, b_k$, $1 < k$. Then there exists an element b_i, say b_k, such that $|b_i| \nless |b_k|$. As proved above, $J = \bar{b}_k + J'$. J' contains the elements $b_1, \cdots, b_{k-1}$, since $(J, \mathbf{s} \leqq |x|) = (J', \mathbf{s} \leqq |x|)$ for $\mathbf{s} \nleqq |b_k|$. Since J' satisfies (a)–(c) as a direct summand of J, and since the elements $b_1, \cdots, b_{k-1}$ form a primitive set of $k - 1 < k$ elements in J', it can be assumed by complete induction that the elements $b_1, \cdots, b_{k-1}$ form a basis of a direct summand of J' and consequently it has been proved that F is a basis of a direct summand of J.

4. Suppose that J satisfies the conditions (b) and (d) and that $b \neq 0$ is an element of J. Then there exists corresponding to every genus $\mathbf{s}$ in $\mathbf{A}(b)$ by (2.4) an element $b'(\mathbf{s})$ such that

$$b \equiv b'(\mathbf{s}) \bmod (J, |x| \nleqq \mathbf{s}), \qquad b'(\mathbf{s}) \equiv 0 \bmod (J, \mathbf{s} \leqq |x|),$$

and there exists by Corollary 8.2 a primitive element $b(\mathbf{s})$ of genus $\mathbf{s}$ in J such that $b'(\mathbf{s}) \equiv b(\mathbf{s}) \bmod (J, \mathbf{s} < |x|)$. The elements $b(\mathbf{s})$ with $\mathbf{s}$ in $\mathbf{A}(b)$ form a primitive set in J. If b' is the sum of this primitive set, then every genus in $\mathbf{A}(b - b')$ is a true multiple of at least one genus in $\mathbf{A}(b)$, since for a genus $\mathbf{s}$ in $\mathbf{A}(b)$, $b - b' \equiv 0 \bmod (J, |x| \nleqq \mathbf{s})$ holds and for a genus $\mathbf{t}$ such that $\mathbf{s} \nleqq \mathbf{t}$ for every $\mathbf{s}$ in $\mathbf{A}(b)$, $b \equiv 0 \bmod (J, |x| \nleqq \mathbf{t})$ holds, and therefore $b - b' \equiv 0 \bmod (J, |x| \nleqq \mathbf{t})$.

Since $\mathbf{A}(b) \leqq \mathbf{A}(J) \leqq \mathbf{B}(J)$, it follows from this result that if the conditions (b) and (d) are satisfied in J and if the chain condition holds in $\mathbf{B}(J)$, every element $\neq 0$ of J is the sum of a primitive set in J.

5. If J satisfies (a)–(d) and if the chain condition holds in $\mathbf{B}(J)$, every element of J is contained in a completely reducible direct summand of J, as has been proved in 3 and 4. Thus it follows from a remark made in 2 that J is separable.

COROLLARY 11.4. *If S is a direct summand of the separable group J and if the chain condition holds in $\mathbf{B}(S)$, then S is separable.*

For J satisfies as a separable group the conditions (a)–(d) of Theorem 11.3 and consequently the direct summand S satisfies these conditions too and is therefore by Theorem 11.3 separable.

THEOREM 11.5. *If J is a separable group, if every group $J(\mathbf{s})^*$ belongs to a class Γ_ν, and if the chain condition holds in $\mathbf{B}(J)$, then J is completely reducible.*

Proof. Every group $J(\mathbf{s})^*$ is by Corollary 8.7, Lemma 11.1 and (2.7) a direct sum of rational groups of genus $\mathbf{s}$. J is therefore by (2.7) and Theorem 10.1 completely reducible.

Remark. Since it can be proved that the group J of Example 9.4 is separable, the chain condition cannot be omitted in Theorem 11.5.

A proof of the separability of the mentioned group J runs as follows.

1. If $v \neq 0$ is an element of J, put $z(v) =$ minimum of the $m(v_i < R_i)_f$ with $v_i \neq 0$.

2. If $v \neq 0$, $z(d) = 1$, let k be the index such that $z(v) = m(v_k < R_k)_f$. Then $v = v' + v''$ with $v'_i = v_i$, $v''_i = 0$ for $i < k$, and $v'_i = 0$, $v''_i = v_i$ for $k \leqq i$. If $\overline{v''}$ is the closed subgroup of J generated by v'', then

$$J = \sum_{i=1}^{k-1} R_i + \overline{v''} + (J, \mathbf{r}_k < |x|),$$

and v is therefore contained in a direct summand of finite rank.

3. If $1 < z(v)$, then $v = v' + v'' + v'''$ with

$$\begin{aligned} v'_i &= v_i, & v''_i &= 0, & v'''_i &= 0 & (i < k), \\ v'_i &= 0, & v''_i &= z(v)z_iw_i, & v'''_i &= z'_iw_i & (k \leqq i), \end{aligned}$$

where $w_i = m(v_i < R_i)_f^{-1}v_i$, if $v_i \neq 0$, $w_i = 0$, if $v_i = 0$, and

$$m(v_i < R_i)_f = z(v)z_i + z_i', \qquad 0 \leqq z_i' < z(v), \qquad \text{if } k \leqq i, v_i \neq 0.$$

Then $z(v''') < z(v)$, for v''' an element of $(J, \mathbf{r}_k < | x |)$ and

$$J = \sum_{i=1}^{k-1} R_i + \overline{v''} + (J, \mathbf{r}_k < | x |).$$

Since the same argument as on J can be applied on $(J, \mathbf{r}_k < | x |)$, by complete induction there exists a completely reducible subgroup V of finite rank and an index i such that v is contained in V and $J = V + (J, \mathbf{r}_i < | x |)$.

4. The separability of J follows now by complete induction with regard to the number of elements contained in the given finite subset. Note furthermore that there exists a continuum of different genera and that therefore the Theorem 11.5 is not a consequence of Theorem 4.7, even if the groups $J(\mathbf{s})^*$ are all countable.

THEOREM 11.6. *The class X of groups is the class of all the separable groups if, and only if, X is the greatest class of groups such that*

(a) *the elements $\neq 0$ of a group J in X are sums of primitive sets in J;*

(b) *every primitive set in a group J in X is a basis of a direct summand of J;*

(c) *if J belongs to X and S is a direct summand of finite rank of J, then J/S belongs to X.*

Proof. The class of all the separable groups satisfies (a)–(c) by Lemma 4.6, Lemma 5.2 and Lemma 5.8. If on the other hand the class X satisfies (a)–(c), all the groups in J are separable, i.e., the class of all the separable groups is the greatest class satisfying (a)–(c).

If n is any finite or infinite cardinal number, C a complete group of rank n, then a group J is isomorphic with a subgroup of C if, and only if, $r(J) \leqq n$. Therefore it is possible to transform Theorem 11.6 into the following criterion which does not use concepts conflicting with the antinomies of the theory of sets.

COROLLARY 11.7. *If n is any finite or infinite cardinal number, C a complete group of rank n and X the greatest class of subgroups of C which satisfies*

(a) *the elements $\neq 0$ of a group J in X are sums of primitive sets in J;*

(b) *every primitive set in a group J in X is a basis of a direct summand of J;*

(c) *if J belongs to X and S is a direct summand of finite rank of J, every subgroup of C (or J) which is isomorphic with J/S belongs to X;*

a group J is isomorphic with a group in X if, and only if, J is separable and $r(J) \leqq n$.

12. **Vector-groups.**

DEFINITIONS 12.1. *If $\phi = (G_1, \cdots, G_v, \cdots)$ is a set of (equal or different) groups, a vector (in ϕ) is a single-valued function (of the indices v) such that the v-th coördinate f_v of the vector f is an element of the group G_v.*

The sum $f + g$ of the two vectors f and g is defined by $(f + g)_v = f_v + g_v$.

Thus the set of all the vectors in ϕ becomes the additive group $V = V(\phi)$ of all the vectors in ϕ.

The group of all the (countable) sequences of integers is an example of a vector-group. Another example is the group discussed as Example 9.4.

If V is a vector-group, then

V' is the subgroup of all those vectors f such that almost every coördinate of f is 0;

V'_v is the subgroup of all those vectors f such that every coördinate but the v-th coördinate is 0;

V''_v is the subgroup of all those vectors f such that the v-th coördinate is 0, and $V = V'_v + V''_v$ for every v, $V' = \sum_v V'_v$.

The following proposition is easily verified:

(12.2) *The group A is isomorphic with a subgroup B of the vector-group $V(\phi)$, satisfying $V' \leqq B \leqq V$ if, and only if, there exist pairs of subgroups A'_v, A''_v of A such that A'_v and G_v are isomorphic; $A = A'_v + A''_v$ for every v; $A'_v \leqq A''_u$ for $v \neq u$; the intersection of the groups A''_v is* 0.[21]

Finally $V' = V(\phi)$ if, and only if, ϕ is finite.[22]

If n is a finite or infinite cardinal number, G any group, ϕ a set consisting of n groups which are isomorphic with G, the notation $V(G, n) = V(\phi)$ will be used.

(12.3) *Suppose that R is a rational group of genus* $\mathbf{s}$.

(a) $\mathbf{s}$ *is an invariant of $V(R, n)$.*

(b) *n is for every rational group R an invariant of $V(R, n)$ if, and only if, n is the only solution of the equation $2^x = 2^n$.*[23]

Proof. Let $\mathbf{s}^0$ be the genus of the infinite part of the numbers of genus $\mathbf{s}$. Then there exist elements of genus $\mathbf{t}$ in $V(R, n)$ if, and only if, $\mathbf{s}^0 \leqq \mathbf{t} \leqq \mathbf{s}$ for infinite n; $\mathbf{s} = \mathbf{t}$ for finite n. This proves (a).

If n is finite, then $r(V(R, n)) = n$ is an invariant of $V(R, n)$.

If n is infinite, then $r(V(R, n)) = 2^n$ is an invariant of $V(R, n)$ and n is conse-

[21] This "approximate decomposition" of the group A into the groups A'_v is exactly the dual of L. Pontrjagin's concept of direct decomposition of the character group of A. But not every approximate decomposition can be realized by a direct decomposition of A. See L. Pontrjagin, *The theory of topological commutative groups*, Ann. of Math., vol. 35 (1934), pp. 361–388.

[22] Though $V(\phi)$ is uniquely determined by the set ϕ, ϕ is not always an invariant of the group $V(\phi)$. If e.g., ϕ consists of 2 groups, isomorphic with the additive group of all the real numbers, and ψ consists of 3 groups, isomorphic with the additive group of all the complex numbers, then $V(\phi)$ and $V(\psi)$ are both complete groups of rank $2^{\aleph_0}$ and therefore isomorphic.

A group which is given as a vector-group is therefore only given by a "representation" of the group and not by invariant properties. This applies also to the topology in the group which is suggested by this representation, since such a topology is generally not a natural topology even in the weak sense that all the proper automorphisms are continuous. Natural topologies in the stronger sense that every proper or improper automorphism is continuous are rather rare. For if in the abelian group A there exists a non-discrete, natural topology in the stronger sense, then A is direct irreducible if A is connected.

[23] n is the only solution of the equation $2^n = 2^x$, if the Cantor continuum hypothesis $2^{\aleph_v} = \aleph_{v+1}$ is satisfied. Without the hypothesis this proposition has not yet been proved. See W. Sierpinski, *L'Hypothèse du Continu.*

quently an invariant of $V(R, n)$, if n is the only solution of $2^x = 2^n$. If w is another solution of this equation, then n and w are infinite. Let W be the additive group of all the rational numbers. Then $V(W, n)$ and $V(W, w)$ are both complete groups of the same rank and therefore isomorphic, i.e., n is not an invariant.[24]

THEOREM 12.4. *Suppose that R is a rational group of genus* $\mathbf{s}$. *Then $V(R, n)$ is completely reducible if, and only if, n is finite or R complete.*

Proof. If n is finite, then $V(R, n)$ is a direct sum of n groups, isomorphic with R, i.e., completely reducible.

If n is infinite and R complete, then $V(R, n)$ is complete and therefore a direct sum of 2^n groups, isomorphic with R, i.e., completely reducible.

Suppose now that n is infinite and R a true subgroup of the additive group of all the rational numbers ($R \neq 0$). If, as may be assumed, the indices of the coördinates are ordinal numbers, denote by W the subgroup of those vectors whose coördinates with infinite index are 0. Then W is a direct summand of $V(R, n)$ and is a $V(R, \aleph_0)$.

Since R is not complete, there exists a prime number p such that $pR < R$. Denote by W_p the subgroup of all those vectors in $(W, \mathbf{s} \leqq |x|)$ such that for every integer i ($\geqq 0$) almost every coördinate is contained in p^iR. Then W_p is a closed subgroup of $(W, \mathbf{s} \leqq |x|)$.

If W' is the subgroup of all those vectors in W such that almost every coördinate is 0, then W' is a direct sum of $\aleph_0$ groups of genus $\mathbf{s}$ and thus W' is countable. $W' \leqq W_p$.

If b is any element in W_p, there exists an element b' in W' such that $b \equiv b'$ mod pW_p. W_p/pW_p is therefore countable. Since all the elements of W_p have in W_p genus $\mathbf{s}$ and since the p-values of their multiplicities are finite, and since finally W_p contains a continuum of elements, this implies that W_p is not completely reducible.

Since $(W, \mathbf{s} < |x|) = 0$ and all the elements $\neq 0$ of $(W, \mathbf{s} \leqq |x|)$ have genus $\mathbf{s}$ in W, and since all the elements $\neq 0$ of W_p have genus $\mathbf{s}$ in W_p, it follows from Corollary 8.9 that $(W, \mathbf{s} \leqq |x|)$ is not completely reducible.

Since all the elements of $(V(R, n), \mathbf{s} \leqq |x|)$ have genus $\mathbf{s}$ in $V(R, n)$ and since $(W, \mathbf{s} \leqq |x|)$ is a direct summand of $(V, \mathbf{s} \leqq |x|)$, it follows that $(V, \mathbf{s} \leqq |x|)$ and consequently that $V(R, n)$ is not completely reducible. Thus the theorem and the following corollary have been proved.

COROLLARY 12.5. *If n is infinite and the rational group R not complete, then $(V(R, n), |R| \leqq |x|)$ is not completely reducible.*

THEOREM 12.6. *If R is a rational group, then $(V(R, n), |R| \leqq |x|)$ is separable.*

Proof. Denote by $\mathbf{s}$ the genus of R and by g a number of genus $\mathbf{s}$. If f is an element of $V(R, n)$ and its v-th coördinate $\neq 0$, then $m(f_v < R) =$

[24] Whether the invariants $|R|$ and 2^n characterize $V(R, n)$ is still unknown.

$\bar{h}(f, v)k(f, v)^{-1}g$, where $\bar{h}$ and k are relatively prime integers, both relatively prime to the infinite part of g. If $f_v = 0$, put $\bar{h}(f, v) = 0$, $k(f, v) = 1$. Since, if $f \neq 0$, $m(f < V)$ is the g.c.d. of the numbers $m(f_v < R)$, f belongs to $(V, \mathbf{s} \leqq |x|)$ if, and only if, the l.c.m. $k(f)$ of all the numbers $k(f, v)$ is finite.

If $f \neq 0$ is an element of $(V, \mathbf{s} \leqq |x|)$, i.e., if $k(f)$ is an ordinary positive integer, put $h(f, v) = \bar{h}(f, v)k(f)k(f, v)^{-1}$. $h(f, v)$ is either 0 or an ordinary positive integer and satisfies $h(f, v)k(f)^{-1} = \bar{h}(f, v)k(f, v)^{-1}$.

If $f \neq 0$ is an element of $(V, \mathbf{s} \leqq |x|)$, denote by $h(f)$ the minimum of the numbers $\bar{h}(f, v) \neq 0$.

1. If $h(f) = 1$, let $\bar{f}$ be the closed subgroup of V generated by f. Since $m(f < J) = k(f)^{-1}g = m(f_u < R)$ for a certain u, it follows that $V = \bar{f} + V''_u$.

2. If $1 < h(f)$, then $f = f' + f''$, where

$$f'_v = z'(v)h(f)h(f, v)^{-1}f_v, \qquad f''_v = z''(v)h(f, v)^{-1}f_v, \qquad (f_v \neq 0),$$

and $f'_v = f''_v = 0$, if $f_v = 0$,

$$h(f, v) = h(f)z'(v) + z''(v), \qquad 0 \leqq z''(v) < h(f).$$

That f'_v and f''_v exist in R is a consequence of the definition of $k(f)$.

Then $1 = h(f')$ and therefore, as proved in 1, $V = \bar{f}' + V''_u$, where u is a suitable index and $\bar{f}'$ the closed subgroup of V generated by f'. Since f'' is an element of V''_u, since $h(f'') < h(f)$ and since V''_u is isomorphic with $V(R, n - 1)$, it follows by complete induction that there exist two subgroups A and B of V such that $V = A + B$, $f \equiv 0 \bmod A$, B consisting of all those vectors b in V such that $b_{v_1} = \cdots = b_{v_q} = 0$ for a given finite set of indices v_i, and A being a direct sum of q rational groups of genus $\mathbf{s}$.

Since B is isomorphic with $V(R, n - q)$, it follows by complete induction that every finite subset of $(V, \mathbf{s} \leqq |x|)$ is contained in a direct summand D of V which is a direct sum of a finite number of rational groups of genus $\mathbf{s}$.

Since this direct summand D is a subgroup of $(V, \mathbf{s} \leqq |x|)$, this implies in particular that $(V, \mathbf{s} \leqq |x|)$ is separable.

By Corollary 4.4 the above result implies the

COROLLARY 12.7. *Every closed subgroup of finite rank of $(V(R, n), |R| \leqq |x|)$ is a direct summand of $V(R, n)$.*

Thus the groups $(V(R, n), |R| \leqq |x|)$ with infinite n and incomplete R are examples of separable groups which are not completely reducible, though all their elements $\neq 0$ have the same genus, and which therefore by Corollary 8.5 do not belong to a class Γ_ν.

COROLLARY 12.8. *Every closed subgroup of $(V(R, n), |R| \leqq |x|)$ is separable and every closed subgroup of $(V(R, n), |R| \leqq |x|)$ which belongs to a class Γ_ν is completely reducible.*

This is a consequence of Corollaries 12.7, 4.4 and 8.7.

If J is a group and $\mathbf{s}$ a genus such that for every closed subgroup S of finite rank all the non-zero elements of J/S have the genus $\mathbf{s}$, it is undecided whether or not J is separable.

THE INSTITUTE FOR ADVANCED STUDY.

SOLUTION OF A PROBLEM OF F. RIESZ ON THE HARMONIC MAJORANTS OF SUBHARMONIC FUNCTIONS

By Tibor Radó

Introduction. Let $u(x, y)$ be a subharmonic function[1] in a domain G. Consider a domain G' comprised in G together with its boundary B'. If $H(x, y)$ is continuous in $G' + B'$ and harmonic in G', and if $H \geqq u$ on B', then $H \geqq u$ in G' also, by the definition of a subharmonic function. If u is continuous, and if the Dirichlet problem is solvable for the region $G' + B'$, then the harmonic function $\bar{h}$ determined by the condition $\bar{h} = u$ on B' is clearly the one which yields the best possible limitation for u on the basis of the fundamental property of subharmonic functions quoted above.

If however u is a general (and therefore possibly discontinuous) subharmonic function, then the situation is less clear. It can be shown (R 2, p. 358) that there exists in G' a *least* harmonic majorant h^* characterized by the following properties. (a) $h^* \geqq u$ in G', (b) if H is harmonic in G' and $H \geqq u$ in G', then $H \geqq h^*$ in G'. But this *least* harmonic majorant did not seem to be the *best* one, as far as usefulness was concerned. At any rate, F. Riesz (R 1, p. 334) reserved the name of best harmonic majorant for a harmonic majorant defined in a different fashion, namely, in terms of the values of u on the boundary B' of G' (see 1.2), while the least harmonic majorant h^* is defined in terms of the values of u in G' alone. *The best harmonic majorant*, in the sense of F. Riesz, will be denoted by $\bar{h}$ and will be referred to by the letters B. H. M. The letters L. H. M. and the notation h^* will refer to *the least harmonic majorant* described above.

F. Riesz stated (R 1, footnote on p. 334) that he established the identity of $\bar{h}$ and h^* in various special cases. Brelot[2] gave an explicit proof in the case when the subdomain G' is bounded by circles. *It is the purpose of this paper to prove the identity of $\bar{h}$ and h^* without any restrictions on G', except for the assumption, implied in the very definition of $\bar{h}$, that the Dirichlet problem is solvable for the region $G' + B'$.*[3]

The proof of this result could be based on the general theorems of F. Riesz

Received October 8, 1936; presented to the American Mathematical Society, December 1936.

[1] See F. Riesz, *Sur les fonctions subharmoniques et leur rapport à la théorie du potentiel*, parts I and II, Acta Mathematica, vol. 48 (1926), pp. 330–343 and vol. 54 (1930), pp. 322–360. These papers will be referred to as R 1 and R 2.

[2] M. Brelot, *Etude des fonctions sousharmoniques au voisinage d'un point*, Actualités scientifiques et industrielles, vol. 139 (1934), p. 18.

[3] In §5 of this paper, we shall give an interpretation of this result which seems to express more adequately its true meaning.

on the representation of subharmonic functions in terms of negative mass distributions. It seemed desirable, however, to present a proof based on much more elementary considerations. The idea of the proof is to push the use of the method of approximation by integral means[4] a little farther than usual and thus to avoid the use of improper integrals which seem to obscure some of the facts which are essential for our purposes. Specifically, our main tool is a trivial modification (formula 2 in 2.2) of the classical formula of Green. Roughly speaking, we apply the formula not to the function itself which is to be studied, but to its integral mean. It seems to the author that this modified formula might prove useful in various other problems, also, whenever it is desirable to avoid improper integrals.

It should be noted that the least harmonic majorant, as defined above, admits of an important interpretation in potential theory. The potential of a negative mass-distribution is a subharmonic function u. The *sweeping-out process*, applied in a domain G', leads to a new potential u^* (see G. C. Evans, *Potentials of positive mass*, part II, Transactions Amer. Math. Soc., vol. 38 (1935), pp. 201–236), and it follows immediately that $u^* = h^*$ in G', where h^* is the least harmonic majorant of u in G'. This remark suggests various problems, similar to the one solved in this paper, which the author plans to discuss elsewhere.

1. **Preliminaries.** 1.1. Let G' be a bounded domain (connected open set). Denote by B' the boundary of G'. We shall say that $G' + B'$ is a Dirichlet region if for every continuous function φ given on B' there exists in $G' + B'$ a continuous function h which is harmonic in G' and which reduces to φ on B'.

1.2. Denote by u a function which is subharmonic in a domain G. Consider a Dirichlet region $G' + B'$ interior to G. As u is upper semi-continuous (R 1, p. 333), we have on B' a sequence of continuous functions φ_k, $k = 1, 2, \cdots$, such that $\varphi_k \searrow u$ on B' (the symbol $\varphi_k \searrow u$ indicates that φ_k is a decreasing sequence). Denote by H_k the solution of the Dirichlet problem, for the region $G' + B'$, with φ_k as the prescribed boundary function. Then (R 1, p. 333) $H_k \searrow \bar{h}$ in G', where $\bar{h}$ is harmonic and $\geqq u$ in G'. This harmonic function $\bar{h}$ is *the best harmonic majorant* (B. H. M.) *in the sense of F. Riesz*[5] of u in G'. We shall list presently some of its properties.

1.3. The B. H. M. depends only upon the values of u on B'. If we use a different sequence $\varphi_k \searrow u$ on B', we obtain the same $\bar{h}$. Also, if u_1, u_2 are subharmonic in G and if $u_1 = u_2$ on B', then u_1 and u_2 have the same B. H. M. in G' (R 1, p. 334).

[4] See R 2, pp. 343 and 345 for historical references concerning the use of this method in the theory of harmonic and subharmonic functions.

[5] Actually, F. Riesz required that φ_k be continuous and $\varphi_k \searrow u$ *in the whole domain* G, while we only require that these conditions be satisfied on B'. The equivalence of the two definitions can be seen immediately by the same reasoning which F. Riesz used to establish the facts stated in 1.3 above. The wording chosen in our text is due to Brelot, loc. cit., p. 17.

1.4. If H is continuous in $G' + B'$, harmonic in G', and $H \geqq u$ in $G' + B'$, then the B. H. M. of u in G' is $\leqq H$ in G' (R 1, p. 334).

1.5. Let $\bar{h}$ be the B. H. M. of u in G', and define in G a function $\bar{u}$ as follows:

$$\bar{u} = \begin{cases} u \text{ in } G - G', \\ \bar{h} \text{ in } G'. \end{cases}$$

Then $\bar{u}$ is subharmonic in G. To see this, take a subregion $G'' + B''$ such that $G' + B' \subset G''$, $G'' + B'' \subset G$. Since u is upper semi-continuous in G, we have in G'' a sequence of continuous functions g_k such that $g_k \searrow u$ in G''. Define in G'' the functions $\bar{g}_k$ as follows:

$$\bar{g}_k = \begin{cases} g_k \text{ in } G'' - G', \\ H_k \text{ in } G', \end{cases}$$

where H_k is the solution of the Dirichlet problem for $G' + B'$ with the boundary condition $H_k = g_k$ on B'. Then by 1.3 we have $\bar{g}_k \searrow \bar{u}$ in G''. Thus $\bar{u}$ is the limit of a decreasing sequence of continuous functions and hence it is upper semi-continuous in G''. As G'' was an arbitrary subdomain, $\bar{u}$ is upper semi-continuous in G. To prove that $\bar{u}$ is subharmonic it must be shown that

$$\bar{u}(x, y) \leqq \frac{1}{2\pi} \int_0^{2\pi} \bar{u}(x + r \cos \varphi, y + r \sin \varphi)\, d\varphi$$

for every point (x, y) in G and for every sufficiently small r. If (x, y) is either in $G - (G' + B')$ or in G', then the inequality is obviously satisfied for small values of r. If (x, y) is on B', we have

$$\begin{aligned} \bar{u}(x, y) = u(x, y) &\leqq \frac{1}{2\pi} \int_0^{2\pi} u(x + r \cos \varphi, y + r \sin \varphi)\, d\varphi \\ &\leqq \frac{1}{2\pi} \int_0^{2\pi} \bar{u}(x + r \cos \varphi, y + r \sin \varphi)\, d\varphi, \end{aligned}$$

since $u \leqq \bar{u}$ in G (see 1.2).

1.6. If u is continuous on B', then (see 1.3) the B. H. M. of u in G' is simply the solution of the Dirichlet problem for $G' + B'$ with the boundary condition $\bar{h} = u$ on B' (it is assumed that $G' + B'$ is a Dirichlet region).

1.7. In G', defined as in 1.2, we have a (necessarily unique) *least harmonic majorant* h^* (L. H. M.) defined as follows: (a) $h^* \geqq u$ in G'; (b) if H is harmonic and $\geqq u$ in G', then $H \geqq h^*$ in G' (see R 2, p. 358).

1.8. Clearly, $h^* \leqq \bar{h}$. As stated in the introduction, we shall prove that $h^* = \bar{h}$. Let us observe that this is trivial if u is continuous. Indeed, take any point (x_0, y_0) on B' and a sequence $(x_n, y_n) \to (x_0, y_0)$ in G'. Since $u \leqq h^* \leqq \bar{h}$, and $u(x_n, y_n) \to u(x_0, y_0) = \bar{h}(x_0, y_0) = \lim \bar{h}(x_n, y_n)$, we have $h^*(x_n, y_n) - \bar{h}(x_n, y_n) \to 0$. That is, the harmonic function $h^* - \bar{h}$ vanishes continuously on the boundary of G', and consequently $h^* = \bar{h}$ in G'.

1.9. The subdomain G' being given as in 1.2, define in G a function u^* as follows:

$$u^* = \begin{cases} u \text{ in } G - G', \\ h^* \text{ in } G', \end{cases}$$

where h^* is the least harmonic majorant of u in G'. *Then u^* is subharmonic in G.*

We first show that u^* is upper semi-continuous in G. Clearly, it is sufficient to verify this property for points on B'. Let (x_0, y_0) be a point on B'. As $u^* \leqq \bar{u}$ (see 1.5 for the definition of $\bar{u}$), we have, for $(x_n, y_n) \to (x_0, y_0)$,

$$\overline{\lim}\, u^*(x_n, y_n) \leqq \overline{\lim}\, \bar{u}(x_n, y_n) \leqq \bar{u}(x_0, y_0) = u(x_0, y_0) = u^*(x_0, y_0).$$

The relation

$$u^*(x, y) \leqq \frac{1}{2\pi} \int_0^{2\pi} u^*(x + r\cos\varphi, y + r\sin\varphi)\, d\varphi$$

is proved exactly as it was proved for $\bar{u}$ in 1.5.

1.10. For $r > 0$ define

$$A_r^{(1)}(x, y) = \frac{1}{r^2\pi} \iint_{\xi^2 + \eta^2 < r^2} u(x + \xi, y + \eta)\, d\xi d\eta$$

(see **R 2**, footnotes on p. 343 and p. 345 for historical references concerning the use of this approximating function in the theory of harmonic and subharmonic functions). Then $A_r^{(1)}(x, y)$ is continuous and subharmonic at those points of G whose distance from the boundary is greater than r. If we apply the same process to $A_r^{(1)}(x, y)$ twice, using the same r, we obtain a function $A_r^{(3)}(x, y)$ which has a number of important properties, some of which will be listed presently (in this section, u always stands for a subharmonic function).

1.11. Take any region $G' + B'$ interior to G. Then for r sufficiently small, $A_r^{(3)}$ is defined and subharmonic in a domain containing $G' + B'$, and has continuous derivatives of the first and second order there. On $G' + B'$, $A_r^{(3)} \searrow u$ for $r \searrow 0$. As a consequence, $\int A_r^{(3)} \to \int u$ for $r \to 0$, where $\int$ stands for any simple or double integral taken over any measurable range in $G' + B'$ (**R 2**, pp. 342–345).

1.12. If $G' + B'$ is interior to G, then (**R 2**, p. 353)

$$\overline{\lim_{r \to 0}} \iint_{G' + B'} \Delta A_r^{(3)}(x, y)\, dx dy < +\infty, \qquad \Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}.$$

1.13. If $G' + B'$ is interior to G, and if u is *harmonic* in G', then we have $A_r^{(3)} = u$ at every point of G' whose distance from B' is larger than $3r$, by the mean-value property of harmonic functions. Consequently for every closed set S in G' we have a $\delta > 0$ such that $\Delta A_r^{(3)} = 0$ on S for $r < \delta$.

1.14. Consider any bounded domain G' with boundary B'. By using subdivisions of the plane into smaller and smaller squares, we can obtain in a familiar fashion an increasing sequence of nested domains G'_n with smooth

boundaries B'_n which approximate $G' + B'$ in the following sense[6] (the properties to be listed are clearly not independent of each other and we state them explicitly only for easier reference).

(a) G'_n contains a prescribed point (x_0, y_0) in G'.

(b) $G'_n + B'_n \subset G'_{n+1}$.

(c) $G'_n + B'_n \subset G'$.

(d) Any point (x, y) of G' is contained in some G'_n.

(e) To every $\epsilon > 0$ there corresponds an $n_0 = n_0(\epsilon)$ with the following property. Denote by S_ϵ the set of those points of G' whose distance from B' is less than ϵ. Then for $n > n_0$ the set S_ϵ contains B'_n.

(f) The measure of G'_n converges to the measure of G', and consequently the measure of $G' - G'_n$ converges to zero.

(g) B'_n consists of a finite number of non-intersecting simple closed curves as smooth as desired. In particular, $G'_n + B'_n$ is a Dirichlet region.

1.15. We shall need the following well-known fact concerning the dependence of the solution of the Dirichlet problem upon the boundary conditions. Suppose the region $G' + B'$ of 1.14 is a Dirichlet region. Denote by F a function which is continuous on $G' + B'$. Let h denote the solution of the Dirichlet problem for $G' + B'$ with the boundary condition $h = F$ on B', and let h_n be the solution of the Dirichlet problem for $G'_n + B'_n$ with the boundary condition $h_n = F$ on B'_n, where $G'_n + B'_n$ is a sequence of approximating regions as described in 1.14. Then $h_n \to h$ in the sense that to every $\eta > 0$ there corresponds a $n_0 = n_0(\eta) > 0$ such that $| h_n - h | < \eta$ in $G'_n + B'_n$ for $n > n_0$. We sketch the proof for the convenience of the reader. Given $\eta > 0$, we have an $\epsilon = \epsilon(\eta) > 0$ such that

$$\left.\begin{aligned} | h(x_2, y_2) - h(x_1, y_1) | &< \frac{\eta}{2} \\ | F(x_2, y_2) - F(x_1, y_1) | &< \frac{\eta}{2} \end{aligned}\right\} \text{ for } [(x_2 - x_1)^2 + (y_2 - y_1)^2]^{\frac{1}{2}} < \epsilon,$$

for every pair of points (x_1, y_1), (x_2, y_2) in $G' + B'$, since F and h are uniformly continuous in $G' + B'$. Take then $n > n_0(\epsilon)$, where $n_0(\epsilon)$ is the quantity defined in condition (e) of 1.14. Consider any point (x, y) on B'_n. By condition (e), we have some point $(\bar{x}, \bar{y})$ on B' such that $[(x - \bar{x})^2 + (y - \bar{y})^2]^{\frac{1}{2}} < \epsilon$. We obtain then

$$\begin{aligned} | h(x, y) - h_n(x, y) | &\leqq | h(x, y) - h(\bar{x}, \bar{y}) | + | h(\bar{x}, \bar{y}) - h_n(x, y) | \\ &= | h(x, y) - h(\bar{x}, \bar{y}) | + | F(\bar{x}, \bar{y}) - F(x, y) | < \eta . \end{aligned}$$

[6] See for instance O. D. Kellogg, *Foundations of Potential Theory*, concerning the familiar facts in 1.14 and 1.15, and in §2.

That is, $|h - h_n| < \eta$ on B'_n for $n > n_0 = n_0(\epsilon(\eta))$. As $h - h_n$ is harmonic in $G'_n + B'_n$, it follows that the same inequality holds in G'_n also.

1.16. The continuity of F was actually used for points of $G' + B'$ close to B'. Hence the conclusion in 1.15 remains valid if we only know that F is continuous in $G' + B' - S$, where S is some closed set in G'.

2. **A remark on the formula of Green.**[6] 2.1. Suppose the functions f, g have continuous derivatives of the first and second orders in a domain G', and that these derivatives remain continuous on the boundary B' of G'. If B' consists of a finite number of sufficiently smooth simple closed curves, then we have the classical formula

$$\iint_{G'} (f\Delta g - g\Delta f)\,dxdy = -\int_{B'} \left(f\frac{\partial g}{\partial n_i} - g\frac{\partial f}{\partial n_i}\right) ds, \tag{1}$$

where n_i refers to the interior normal with respect to G'.

2.2. Let there be given in a domain G a function v with continuous derivatives of the first and second orders. Consider a region $G' + B'$, interior to G, such that B' consists of a finite number of simple closed smooth curves. In G', take a circle $C(x_0, y_0; r)$, with centre (x_0, y_0) and radius r, comprised in G' together with its interior. Put

$$v^{(r)}(x_0, y_0) = \frac{1}{2\pi}\int_0^{2\pi} v(x + r\cos\varphi, y + r\sin\varphi)\,d\varphi,$$

$$l(x, y) = -\log[(x - x_0)^2 + (y - y_0)^2]^{\frac{1}{2}}, \qquad (x, y) \neq (x_0, y_0),$$

$$l_r(x, y) = \begin{cases} l(x, y) \text{ for } [(x - x_0)^2 + (y - y_0)^2]^{\frac{1}{2}} \geqq r, \\ -\log r \text{ for } [(x - x_0)^2 + (y - y_0)^2]^{\frac{1}{2}} \leqq r. \end{cases}$$

Denote by h the solution of the Dirichlet problem for $G' + B'$ with the boundary condition $h = v$ on B'. Then (see Kellogg, loc. cit., p. 237, footnote)

$$h(x_0, y_0) = \frac{1}{2\pi}\int_{B'} v\left(\frac{\partial l}{\partial n_i} - \frac{\partial H}{\partial n_i}\right) ds,$$

where H denotes the solution of the Dirichlet problem for $G' + B'$ with the boundary condition $H = l$ on B' (that is, $l - H$ is Green's function for $G' + B'$ with pole at (x_0, y_0)). If B' is smooth, then the necessary derivatives of H will remain continuous on B'.[7]

2.3. We apply now formula (1) *first* to the region bounded by B' and $C(x_0, y_0; r)$ for $f = v, g = l$, *second* to the circular disc bounded by $C(x_0, y_0; r)$ for $f = v, g = -\log r$ and *third* to $G' + B'$ for $f = v, g = H$. Combining the resulting equations with the expression for $h(x_0, y_0)$ in 2.2, we obtain the formula

[7] The functions l, l_r, H depend also upon the choice of (x_0, y_0). However, this point will be kept fixed, and there is no need for notations like $H(x, y; x_0, y_0)$ etc.

(2) $$v^{(r)}(x_0, y_0) = -\frac{1}{2\pi} \iint_{G'} [l_r(x, y) - H(x, y)] \Delta v(x, y)\, dx dy + h(x_0, y_0).$$

2.4. Actually, (2) can be obtained from the classical expression for v in terms of Green's function by an integration under the integral sign. Conversely, that classical formula can be obtained from (2) by the passage to the limit $r \to 0$. But our idea is to keep r finite in (2), so that we have to deal only with continuous functions. We shall need a slight extension of (2) which we shall consider presently.

2.5. If v has continuous derivatives of the first and second orders in G, then formula (2) holds for every Dirichlet subregion $G' + B'$.

This may be seen as follows. Approximate $G' + B'$ by regions $G'_n + B'_n$ as described in 1.14. Observe first that the functions h, H, as defined in 2.2, actually exist for every Dirichlet subregion $G' + B'$. Denote by h_n, H_n the functions corresponding to $G'_n + B'_n$. We have then by 2.2

(3) $$v^{(r)}(x_0, y_0) = -\frac{1}{2\pi} \iint_{G'_n} [l_r(x, y) - H_n(x, y)] \Delta v(x, y)\, dx dy + h_n(x_0, y_0).$$

We have $h_n(x_0, y_0) \to h(x_0, y_0)$ by 1.15. Next we consider

$$\iint_{G'} l_r(x, y) \Delta v(x, y)\, dx dy = \iint_{G'_n} + \iint_{G'-G'_n}.$$

As l_r and Δv are continuous and therefore bounded in $G' + B'$, we have

$$\iint_{G'-G'_n} \to 0,$$

since the measure of $G' - G'_n$ converges to zero. Hence

$$\iint_{G'_n} l_r(x, y) \Delta v(x, y)\, dx dy \to \iint_{G'} l_r(x, y) \Delta v(x, y)\, dx dy.$$

Consider finally

(4) $$\begin{aligned} \iint_{G'} H(x, y) \Delta v(x, y)\, dx dy &= \iint_{G'_n} H_n(x, y) \Delta v(x, y)\, dx dy \\ &+ \iint_{G'_n} [H(x, y) - H_n(x, y)] \Delta v(x, y)\, dx dy \\ &+ \iint_{G'-G'_n} H(x, y) \Delta v(x, y)\, dx dy = I_n^{(1)} + I_n^{(2)} + I_n^{(3)}. \end{aligned}$$

We have $I_n^{(3)} \to 0$ because the integrand is bounded and the measure of $G' - G'_n$ converges to zero. Take now any $\eta > 0$. By 1.16, we shall have for n larger than some $n_0 = n_0(\eta)$

$$| H - H_n | < \eta \text{ in } G'_n + B'_n.$$

Hence, for large n,

$$| I_n^{(2)} | < \eta \iint_{G'_n} | \Delta v(x, y) | \, dxdy \leqq \eta \iint_{G'} | \Delta v(x, y) | \, dxdy.$$

Thus $I_n^{(2)} \to 0$. It follows then from (4) that

$$\iint_{G'_n} H_n(x, y) \Delta v(x, y) \, dxdy \to \iint_{G'} H(x, y) \Delta v(x, y) \, dxdy.$$

Summing up, formula (3) yields, for $n \to \infty$, formula (2) for the most general Dirichlet region $G' + B' \subset G$. The assumptions concerning the smoothness of v could be generalized in an obvious way, but this is immaterial for our purposes.

2.6. If the function v of 2.5 is subharmonic in G, then the function h in formula (2) is the B. H. M. of v in G' (observe that v is continuous by assumption, and compare 2.2 and 1.6).

3. **A lemma.** 3.1. LEMMA. *Let u be subharmonic in a domain G. Denote by $G' + B'$ a Dirichlet region interior to G, and suppose that u is harmonic in G'. Then $u = \bar{h}$ in G', where $\bar{h}$ is the B. H. M. of u in G'.*

3.2. To prove this lemma, consider the sequence of approximating functions

$$u_n(x, y) = A_{1/n}^{(3)} (x, y) \qquad n > N$$

(see 1.10), where N is large enough so that for $n > N$ the function u_n is defined in some domain which contains $G' + B'$ in its interior. Denote by $\bar{h}_n$ the B. H. M. of u_n in G'. Since $u_n \searrow u$, and since u_n is continuous, we have (see 1.2, 1.3)

$$\bar{h}_n \to \bar{h} \text{ in } G', \tag{5}$$

where $\bar{h}$ is the B. H. M. of u in G'. Also (see 1.12)

$$0 \leqq \iint_{G'+B'} \Delta u_n(x, y) \, dxdy < M, \tag{6}$$

where M is some finite constant. In G', take a point (x_0, y_0) and a circle $C(x_0, y_0; r)$ with centre (x_0, y_0) and radius r, such that $C(x_0, y_0; r)$ is comprised in G' together with its interior. On account of 2.5, 2.6 we have then (since u_n has continuous derivatives of the first and second order)

$$(7)\quad \begin{aligned} u_n^{(r)}(x_0, y_0) &= -\frac{1}{2\pi}\iint_{G'} [l_r(x, y) - H(x, y)]\Delta u_n(x, y)\,dxdy + \bar{h}_n(x_0, y_0) \\ &= \frac{1}{2\pi} I_n + \bar{h}_n(x_0, y_0). \end{aligned}$$

3.3. We write

$$(8)\quad I_n = \iint_{G'} [H(x, y) - l_r(x, y)]\Delta u_n(x, y)\,dxdy = \iint_S + \iint_{G'-S},$$

where S is a closed set interior to G'. Observe that $l_r = l = H$ on B' (see 2.2). That is, $H - l_r = 0$ on B'. As l_r, H are continuous in $G' + B'$, we have for every $\epsilon > 0$ a $\delta > 0$ such that $| H(x, y) - l_r(x, y) | < \epsilon$ for every point (x, y) in G' whose distance from B' is less than δ. Choose the closed set S of formula (8) as the set of those points of G' whose distances from B' are $\geqq\delta$. In formula (8) we have then, by (6) in 3.2,

$$(9)\quad \left|\iint_{G'-S}\right| \leqq \epsilon \iint_{G'-S} \Delta u_n(x, y)\,dxdy \leqq \epsilon \iint_{G'+B'} \Delta u_n(x, y)\,dxdy < \epsilon M.^{3}$$

On the other hand, for n large enough we have $\Delta u_n = 0$ on S (see 1.13). Hence in formula (8) we have $\iint_S = 0$ for large n. That is,

$$\overline{\lim}\, | I_n | \leqq \epsilon M.$$

As ϵ is arbitrary and M is fixed, it follows that $I_n \to 0$.

3.4. By 1.11, 3.3 and by formula (5) it follows from formula (7) that $u^{(r)}(x_0, y_0) = \bar{h}(x_0, y_0)$. But u is harmonic in G' and therefore $u^{(r)}(x_0, y_0) = u(x_0, y_0)$. Thus $u(x_0, y_0) = \bar{h}(x_0, y_0)$, and the lemma is proved.

4. **The theorem.** THEOREM. *Let u be subharmonic in a domain G. Let $G' + B'$ be a Dirichlet region interior to G. Denote by $\bar{h}$ and h^* the best harmonic majorant and the least harmonic majorant of u in G' (see 1.2 and 1.7). Then $\bar{h} = h^*$.*

Proof. Define

$$u^* = \begin{cases} u \text{ in } G - G', \\ h^* \text{ in } G'. \end{cases}$$

Then u^* is subharmonic in G (see 1.9). Denote by $\bar{h}^*$ the B. H. M. of u^* in G'. Since u and u^* are both subharmonic in G and $u = u^*$ on B', we have (see 1.3) $\bar{h}^* = \bar{h}$ in G'. Since u^* is harmonic in G', we have (see 3.1) $u^* = \bar{h}^*$ in G'. But by definition $u^* = h^*$ in G'. Hence $h^* = \bar{h}$, as stated in the theorem.

[3] Observe that u_n is subharmonic and therefore $\Delta u_n \geqq 0$ (R 1, p. 335).

5. Conclusion. 5.1. Let u be subharmonic in a domain G. Denote by $G' + B'$ a region interior to G, and by h' a harmonic function in G'. Define in G a function u' as follows:

$$u' = \begin{cases} u \text{ in } G - G', \\ h' \text{ in } G'. \end{cases}$$

If u' is subharmonic in G, we shall say that the harmonic function h' is *admissible* for u in G'.

5.2. We have then the following

THEOREM. *If u is subharmonic in a domain G, and if $G' + B'$ is a Dirichlet region interior to G, then there exists in G' exactly one harmonic function which is admissible in G' for u in the sense of* 5.1.

Indeed, denote by $\bar{h}$ the B. H. M. of u in G'. Then $\bar{h}$ is admissible for u in G', by 1.5. Let h' be any harmonic function which is admissible for u in G', and consider the corresponding function u', as defined in 5.1. Let $\bar{h}'$ be the B. H. M. of u' in G'. Since u and u' are subharmonic in G and $u = u'$ on B', we have $\bar{h} = \bar{h}'$ in G', by 1.3. Since u' is harmonic in G', we have $\bar{h}' = u'$ in G' by 3.1. By definition $u' = h'$ in G'. Hence $h' = \bar{h}$ in G', and the theorem is proved.

5.3. This theorem implies that the best harmonic majorant can be characterized as follows: If $G' + B'$ is a Dirichlet region interior to G, *then, as is obvious from* 1.5 *and* 5.2, *the best harmonic majorant of u in G' can be defined as the unique harmonic function $\bar{h}$ with the property that the function $\bar{u}$, equal to u in $G - G'$ and to $\bar{h}$ in G', is again subharmonic in G.*

THE OHIO STATE UNIVERSITY.

SUMMABILITY OF DOUBLE FOURIER SERIES

By J. J. Gergen

Part I

1.1. **Introduction.** We consider in Part I the M. Riesz and the Cesàro sums of a simple series

$$\sum_{p=0}^{\infty} a_p, \tag{1.11}$$

that is, the sums

$$\sigma_\alpha(x) = \sum_{p<x} (x-p)^\alpha a_p, \qquad S_\alpha(m) = \sum_{p=0}^{m} A_{m-p}^\alpha a_p,$$

where, in the Cesàro sums, A_p^α is the coefficient of x^p in the expansion

$$(1-x)^{-\alpha-1} = 1 + \sum_{p=1}^{\infty} A_p^\alpha x^p.$$

Riesz's theorem that for $0 \leqq \alpha$ the existence of either of the limits

$$\lim_{m\to\infty} S_\alpha(m)/A_m^\alpha = L, \qquad \lim_{x\to\infty} \sigma_\alpha(x)/x^\alpha = L$$

implies that of the other is of course well known.[1] In Theorems I and II below we give some formulas expressing in a simple manner each of the sums in terms of the other. On the basis of these theorems Riesz's result can readily be obtained. The derivations do not in general follow Riesz's procedure; and as regards conciseness there seems to be some advantage in following this second method of approach. Theorems I and II can also be used to advantage to simplify the extensions to two variables of Riesz's theorem given by Dr. S. B. Littauer and the author.[2] In Part II of the present paper we shall apply these theorems in connection with double Fourier series.

Theorem I. *Let $0 < \alpha$. Let k be the integral part of α. Let $f(x)$ be continuous with its derivatives $f', f'', \cdots, f^{(k+2)}$ for $0 \leqq x$, and satisfy*

$$\begin{aligned} f(0) = f'(0) = \cdots = f^{(k+1)}(0) &= 0, \\ f(x) = \Gamma(\alpha+x)/\{\Gamma(x)\Gamma^2(\alpha+1)\} \ &\textit{for } 1 \leqq x. \end{aligned} \tag{1.12}$$

Received December 26, 1936; presented to the American Mathematical Society, April 11, 1936.

[1] For a presentation of Riesz's proof see Hobson, *Theory of Functions of a Real Variable*, vol. 2, 1926, pp. 90–98.

[2] J. J. Gergen and S. B. Littauer, *Continuity and summability for double Fourier series*, Transactions of the American Mathematical Society, vol. 38 (1935), pp. 401–435. This paper will be denoted by A.

Then the fractional derivative $U(x)$,

$$U(x) = f^{(\alpha+1)}(x) = 1/\Gamma(k+1-\alpha) \int_0^x (x-t)^{k-\alpha} f^{(k+2)}(t)dt,$$

of order $\alpha + 1$ of f, is continuous for $0 < x$, and we have

$$S_\alpha(m) = \int_0^{m+1} U(m+1-t)\sigma_\alpha(t)dt \qquad \text{for } m = 0, 1, \cdots, \tag{1.13}$$

$$U(x) = O(x^{-2}), \qquad x \to \infty, \tag{1.14}$$

$$U(x) = O(x^{k+1-\alpha}), \qquad x \to +0. \tag{1.15}$$

THEOREM II. *Let $0 < \alpha$. Let*

$$\varphi_0(x) = \Gamma(\alpha+1) \sum_{p<x} A_p^{-\alpha-2}. \tag{1.16}$$

Then the fractional integral $V(x)$,

$$V(x) = \varphi_\alpha(x) = 1/\Gamma(\alpha) \int_0^x (x-t)^{\alpha-1} \varphi_0(t)dt,$$

of order α of φ_0, is continuous for $0 < x$, and we have

$$\sigma_\alpha(x) = \sum_{p<x} V(x-p)S_\alpha(p) \qquad \text{for } 0 < x, \tag{1.17}$$

$$V(x) = O(x^{-2}), \qquad x \to \infty, \tag{1.18}$$

$$V(x) = O(x^{\alpha}), \qquad x \to +0. \tag{1.19}$$

2.1. **Lemma for Theorems I and II.** Both Theorems I and II depend on the following lemma. To simplify the writing we shall suppose in what follows that x, y are positive numbers, and that m, n, p, q are integers, positive or 0. We denote by M a number independent of m, n, p, q, x, y over the range

$$0 < x, \quad 0 < y; \quad 0 \leqq m, \quad 0 \leqq n, \quad 0 \leqq p, \quad 0 \leqq q,$$

or that part of this range specified. The symbols o, O refer to the behavior of the function in question in the neighborhood of ∞ or (∞, ∞).

LEMMA 1. *Let $0 < \delta < 1$. Let $F_0(x)$ be integrable over every finite interval $0 < x < x_0$ and satisfy*

$$F_p(x) = x^{p-\delta-1}/\Gamma(p-\delta) + O(x^{p-\delta-2})$$

for $p = 0, 1, 2$. Then

$$F_\delta(x) = O(x^{-2}).$$

We have for each fixed b, $0 < b < 1$,

$$\Gamma(\delta)F_\delta(x) = x^\delta \left[\int_0^b + \int_b^1\right] (1-t)^{\delta-1} F_0(tx)dt$$

$$= x^{\delta}\left\{(1-b)^{\delta-1}x^{-1}F_1(bx) + (\delta-1)(1-b)^{\delta-2}x^{-2}F_2(bx)\right.$$

$$\left. + (\delta-1)(\delta-2)x^{-2}\int_0^b (1-t)^{\delta-3}F_2(tx)dt + \int_b^1 (1-t)^{\delta-1}F_0(tx)dt\right\}$$

$$= E(b)x^{-1} + O(x^{-2}),$$

where

$$\Gamma(1-\delta)E(b) = (1-b)^{\delta-1}b^{-\delta} - (1-b)^{\delta-2}b^{1-\delta}$$

$$- (\delta-2)\int_0^b (1-t)^{\delta-3}t^{1-\delta}dt - \delta\int_b^1 (1-t)^{\delta-1}t^{-\delta-1}dt.$$

Now, differentiating, we see that E' vanishes for $0 < b < 1$. Hence, using the expansion

$$\frac{1}{z-x} = \frac{1}{z} + \frac{x}{z}\frac{1}{z+1} + \frac{x}{z}\frac{x+1}{z+1}\frac{1}{z+2} + \cdots$$

with $z = 1$, $x = 1-\delta$, we have

$$\Gamma(1-\delta)E(b) = \lim_{b\to+0}\left\{(1-b)^{\delta-1}b^{-\delta} - \delta\int_b^1 (1-t)^{\delta-1}t^{-\delta-1}dt\right\}$$

$$= 1 - \delta\int_0^1 \{(1-t)^{\delta-1} - 1\}t^{-\delta-1}dt$$

$$= 1 - \delta\{1 + (1-\delta)/2! + (1-\delta)(2-\delta)/3! + \cdots\} = 0.$$

The lemma follows.

3.1. **Proof of Theorem** I. The continuity of $f^{(k+2)}$ implies the continuity of $f^{(\alpha+1)}$ and the truth of (1.15). On the other hand, U satisfies

$$f(x) = 1/\Gamma(\alpha+1)\int_0^x (x-t)^{\alpha}U(t)dt$$

because of the first half of (1.12). Hence

$$\int_0^{m+1} \sigma_\alpha(t)U(m+1-t)dt = \sum_{p=0}^{m} a_p \int_0^{m+1-p} (m+1-p-u)^{\alpha}U(u)du$$

$$= \Gamma(\alpha+1)\sum_{p=0}^{m} f(m+1-p)a_p = S_\alpha(m).$$

Accordingly, it is enough to prove that (1.14) holds.

If α is an integer, then f reduces to a polynomial of degree α for $1 \leqq x$. Hence

$$U(x) = f^{(k+1)}(x) = 0 \qquad \text{for } 1 < x.$$

This gives (1.14) in this case.

Suppose then that α is not an integer, that $\alpha = k + 1 - \delta$, where $0 < \delta < 1$. From Stirling's formula[3]

$$\log \Gamma(x) = (x - \tfrac{1}{2}) \log x - x + \tfrac{1}{2} \log 2\pi + P(0)/x - \int_0^\infty P(t)/(t + x)^2 dt,$$

$$P(t) = \sum_{n=1}^{\infty} \{\cos 2n\pi t\}/\{2n^2\pi^2\},$$

it follows readily that

$$\Gamma^2(\alpha + 1)f(x) = x^\alpha e^{H(x)},$$

where

$$D_x^{(p)} H = O(x^{-p-1}).$$

Hence,

$$\Gamma(\alpha + 1)D^{(p)}f = x^{\alpha-p}/\Gamma(\alpha + 1 - p) + O(x^{\alpha-p-1}).$$

Taking now in Lemma 1

$$F_0(x) = \Gamma(\alpha + 1)f^{(k+2)}(x),$$

we see that F_δ coincides with $\Gamma(\alpha + 1)U$ and that

$$F_p(x) = \Gamma(\alpha + 1)D^{(k+2-p)}f(x) = x^{p-\delta-1}/\Gamma(p - \delta) + O(x^{p-\delta-2})$$

for $p = 0, 1, 2$. Thus (1.14) holds, and this completes the proof.

4.1. **Proof of Theorem** II. We have

$$\varphi_\alpha(x) = \sum_{p<x} A_p^{-\alpha-2} (x - p)^\alpha. \tag{4.11}$$

The continuity of V and the truth of (1.19) then follows. In addition,

$$\sigma_\alpha(x) = \sum_{p<x} (x - p)^\alpha \sum_{q=0}^{p} A_{p-q}^{-\alpha-2} S_\alpha(q) = \sum_{p<x} V(x - p) S_\alpha(p).$$

It remains then to consider (1.18).

Let $\Delta_q x^n$ denote the q-th difference

$$\Delta_q x^n = \sum_{p=0}^{q} A_p^{-q-1}(x - p)^n. \tag{4.12}$$

Then, for $1 \leqq q \leqq n$,

$$\Delta_q x^n = n!/(n - q)! \int_{x-1}^{x} dt_1 \int_{t_1-1}^{t_1} dt_2 \cdots \int_{t_{q-1}-1}^{t_{q-1}} t_q^{n-q}\, dt_q.$$

Hence

$$|\Delta_q x^n| \leqq n!\, x^{n-q} \qquad \text{for } 1 \leqq q \leqq x, \quad q \leqq n,$$

[3] See, for example, Bieberbach, *Lehrbuch der Funktionentheorie*, vol. 1, 1923, pp. 297–308.

$$\Delta_n x^n = n!, \qquad \Delta_{n+1} x^n = 0.$$

Now if α is an integer, we have

$$V(x) = \Delta_{\alpha+1} x^\alpha = 0$$

for $\alpha + 1 < x$; and this is (1.18) in this case.

Suppose then that α is not an integer, that $\alpha = k + \delta$, where $0 < \delta < 1$. Denoting by μ the largest integer less than x, we have

$$\varphi_0(x)/\Gamma(\alpha + 1) = A_\mu^{-\alpha-1} = x^{-\alpha-1}/\Gamma(-\alpha) + O(x^{-\alpha-2}).$$

In addition, for each fixed positive n,

$$\begin{aligned} \varphi_n(x)/\Gamma(\alpha + 1) &= A_{\mu-n}^{-\alpha-1+n} + 1/n! \sum_{q=0}^{n-1} A_{\mu-q}^{-\alpha-1+q} \Delta_q(x - \mu + q)^n \\ &= x^{n-\alpha-1}/\Gamma(n - \alpha) + O\left(x^{n-\alpha-2} + \sum_{q=0}^{n-1} x^{-\alpha-1+q}\right) \\ &= x^{n-\alpha-1}/\Gamma(n - \alpha) + O(x^{n-\alpha-2}). \end{aligned}$$

Taking

$$F_0(x) = \varphi_k(x)/\Gamma(\alpha + 1),$$

we see that F_δ coincides with $\varphi_\alpha/\Gamma(\alpha + 1)$, and that

$$F_p(x) = \varphi_{k+p}(x)/\Gamma(\alpha + 1) = x^{p-\delta-1}/\Gamma(p - \delta) + O(x^{p-\delta-2})$$

for $p = 0, 1, 2$. The lemma is then applicable and the theorem follows.

Part II

5.1. **Introduction.** We consider here a function $f(u, v)$ integrable L over $(0, 0; 2\pi, 2\pi)$ and with period 2π in each variable. We shall suppose in addition that f is even in each variable, and shall consider only the Fourier series of f at the origin, that is, the series

$$\sum_{m,n=0}^{\infty} a_{m,n}, \tag{5.11}$$

where

$$a_{m,n} = \lambda_m \lambda_n \int_0^\pi \cos mu \, du \int_0^\pi \cos nv \, f(u, v) \, dv,$$

$$\lambda_0 = 1/\pi, \qquad \lambda_1 = \lambda_2 = \cdots = 2\lambda_0.$$

These simplifications involve no loss in generality. Our object is to complete some of the results on Cesàro and Riesz summability of the series (5.11) given in A. We are interested in particular in the question whether the rôles of summability and continuity can be interchanged in the following Theorem VI of A.[4]

[4] This question is raised on p. 415.

THEOREM VI. *Suppose that*

$$0 \leqq \alpha < a - 2, \qquad 0 \leqq \beta < b - 2$$

and that the series (5.11) *is summable either* $(C; \alpha, \beta)$ *or* $(R; \alpha, \beta)$ *to sum* s. *Then, if* f *is bounded* $(C; a, b)$ *in the square* $(+0, +0; \delta, \delta)$ *for some* $\delta > 0$, *it follows that* f *is continuous* $(C; a, b)$ *with limit* s.[5]

5.2. In connection with Theorem VI it might be noted here, incidentally, that it carries with it the following permanency theorem.

THEOREM III. *Suppose that* a, b, α, β *are non-negative numbers, that* f *is continuous* $(C; a, b)$ *with limit* s, *and that the series* (5.11) *is summable either* $(R; \alpha, \beta)$ *or* $(C; \alpha, \beta)$ *to sum* S; *then*

$$s = S. \tag{5.21}$$

The proof is immediate. We set

$$\xi = \alpha + a + 3, \qquad \eta = \beta + b + 3.$$

Then, since f is continuous $(C; a, b)$ with limit s, it is continuous $(C; \xi, \eta)$ with limit s.[6] On the other hand, since f is continuous $(C; \xi, \eta)$, it is bounded $(C; \xi, \eta)$ in some square $(+0, +0; \delta, \delta)$, $0 < \delta$. Thus, making use of our summability hypothesis and Theorem VI, we see that f is continuous $(C; \xi, \eta)$ with limit S. We conclude that (5.21) holds.

5.3. We turn now to the question on Theorem VI. The result that one might expect would be that continuity $(C; a, b)$ of f, plus ultimate boundedness, either $(C; \alpha, \beta)$ or $(R; \alpha, \beta)$, of the series (5.11), implies summability of the same type and order of the series for α, β sufficiently large. Peculiarly enough, this is false. This we shall show in paragraphs 6.1 through 7.1, the result of which we state in the form of

THEOREM IV. *Corresponding to every set of non-negative numbers* a, b, α, β *there can be constructed a function* $f(u, v)$ *integrable* L *over* $(0, 0; \pi, \pi)$, *even and periodic with period* 2π *in each variable, and continuous* $(C; a, b)$, *for which the series* (5.11) *is ultimately bounded both* $(R; \alpha, \beta)$ *and* $(C; \alpha, \beta)$, *but is summable neither* $(R; \alpha, \beta)$ *nor* $(C; \alpha, \beta)$.

Actually we shall prove somewhat more than is stated. The function f we construct will vanish on $(0, 0; \frac{1}{2}\pi, \pi)$. In addition, the series (5.11) will not only be ultimately bounded but bounded both $(R; \alpha, \beta)$ and $(C; \alpha, \beta)$.

5.4. The natural analogue of Theorem VI being false, one might ask if it is possible to obtain any result of the same general character. In our final theorem, the proof of which is in paragraphs (8.1) through (9.1), by requiring ultimate boundedness of two different orders, we obtain a result of this kind.

[5] The definitions for R and C summability can be found on pp. 401–402 of A, those concerning the continuity and boundedness of f, on p. 413.

[6] This follows directly from Theorem III, p. 413 of A.

THEOREM V. *Suppose that*

$$(5.41)\quad 0 \leqq a < \xi, \qquad 0 \leqq \alpha < \xi - 1, \qquad 0 \leqq b < \eta, \qquad 0 \leqq \beta < \eta - 1$$

and that f is continuous $(C; a, b)$ *with limit s. Then, if the series* (5.11) *is ultimately bounded both* $(R; \alpha, \eta)$ *and* $(R; \xi, \beta)$, *it is summable* $(R; \xi, \eta)$ *to sum s. In addition, the corresponding result with C in place of R holds.*

6.1. Lemmas for Theorem IV.

LEMMA 2. *For every m there exists an even, bounded and measurable function* $I_m(u)$, *with period* 2π, *which vanishes on* $(0, \frac{1}{2}\pi)$, *and whose Fourier series,*

$$I_m(u) \sim \sum_{p=0}^{\infty} C_{p,m} \cos pu,$$

satisfies

$$C_{0,m} = C_{1,m} = \cdots = C_{m,m} = 0, \qquad C_{m+1,m} \neq 0.$$

To prove the lemma it is enough to show that, m being fixed, we can choose constants $b_0, b_1, \cdots, b_{m+1}$ so that

$$(6.11)\qquad \sum_{q=0}^{m+1} A_{p,q} b_q = 0 \text{ or } 1,$$

according as $p \leqq m$ or $p = m + 1$, where

$$A_{p,q} = \int_{\frac{1}{2}\pi}^{\pi} \cos pu \cos qu \, du.$$

For, with the b's so chosen, the even, periodic function I_m which vanishes on $(0, \frac{1}{2}\pi)$ and coincides with

$$\sum_{q=0}^{m+1} b_q \cos qu$$

for $\frac{1}{2}\pi < u \leqq \pi$ has the desired properties.

Now a short calculation shows that $A_{p,q}$ is rational for $p \neq q$ and has the value $1/(2\lambda_p)$ for $p = q$. Thus the determinant Δ of the A's can be written in the form

$$\Delta = \pi^{m+2}/2^{2m+3} + r_1\pi^{m+1} + \cdots + r_{m+2},$$

where the r's are rational. Since π is not algebraic, it follows that Δ is not 0. Accordingly the equations (6.11) admit a solution, and the lemma follows.

6.2. LEMMA 3. *Let*

$$(6.21)\qquad 0 < \cdots < e_1 < e_0 = \pi; \qquad e_p/e_{p+1} \to \infty, \qquad p \to \infty.$$

Let $\varphi_p(v)$ *be the even, periodic function which coincides with* v^{-1} *for* $e_{p+1} \leqq v < e_p$ *and equals* 0 *elsewhere on* $(0, \pi)$. *Let* $R_\beta(y; \varphi_p)$ *be the Riesz mean,*

$$R_\beta(y; \varphi_p) = \sum_{n<y} (1 - y/n)^\beta \lambda_n \int_0^{\pi} \cos nv \, \varphi_p(v) \, dv,$$

of order β, $0 \leqq \beta$, of the Fourier series of φ_p at the origin. Then

$$|R_\beta(y; \varphi_p)| < M/e_{p+1}^{\frac{1}{2}}, \tag{6.22}$$

and, for every positive ξ,[7]

$$e_{p+1}^{\frac{1}{2}} R_\beta(\xi/e_{p+1}; \varphi_p) \to 2\,\xi^{\frac{1}{2}}/\pi \int_\xi^\infty \gamma_{\beta+1}(v)/v^{\frac{1}{2}}\,dv, \qquad p \to \infty. \tag{6.23}$$

Suppose first that $0 < \beta$. We have

$$\begin{aligned} R_\beta(y; \varphi_p) &= 2\,y/\pi \int_0^\infty \gamma_{\beta+1}(yv)\,\varphi_p(v)\,dv \\ &= 2\,y/\pi \left[\int_{e_{p+1}}^{e_p} + \int_\pi^\infty\right] \gamma_{\beta+1}\,\varphi_p\,dv \\ &= F_1(y) + F_2(y), \end{aligned} \tag{6.24}$$

say.

Applying the second mean value theorem, we get

$$\begin{aligned} |F_1| &= \left| 2\,y/(\pi\,e_{p+1}^{\frac{1}{2}}) \int_{e_{p+1}}^{\eta} \gamma_{\beta+1}(yv)\,dv \right| \qquad (e_{p+1} \leqq \eta \leqq e_p) \\ &< M/e_{p+1}^{\frac{1}{2}} \int_0^\infty |\gamma_{\beta+1}(v)|\,dv < M/e_{p+1}^{\frac{1}{2}}. \end{aligned}$$

In addition, since $e_p/e_{p+1} \to \infty$, we have

$$e_{p+1}^{\frac{1}{2}} F_1(\xi/e_{p+1}) = 2\,\xi^{\frac{1}{2}}/\pi \int_\xi^{\xi e_p/e_{p+1}} \gamma_{\beta+1}(v)/v^{\frac{1}{2}}\,dv \to 2\,\xi^{\frac{1}{2}}/\pi \int^\infty \gamma_{\beta+1}(v)/v^{\frac{1}{2}}\,dv.$$

On the other hand, for $1 \leqq y$,

$$\begin{aligned} |F_2| &< My \int_\pi^\infty |\gamma_{\beta+1}(yv)|\,\varphi_p(v)\,dv \\ &< M(y^{-\beta} + y^{-1}) \int_\pi^\infty (v^{-\beta-1} + v^{-2})\,\varphi_p(v)\,dv < M. \end{aligned}$$

We conclude that (6.22) holds for $1 \leqq y$, and that (6.23) holds. Since $R_\beta(y; \varphi_p) = R_\beta\,(1; \varphi_p)$ for $y < 1$, the lemma follows for $0 < \beta$.

The proof for the case $\beta = 0$ is similar and can be omitted. In this case the formula

$$\begin{aligned} R_0(y; \varphi_p) &= 2\,n/\pi \int_{e_{p+1}}^{e_p} \gamma_1(nv)/v^{\frac{1}{2}}dv \\ &\quad + 1/\pi \int_{e_{p+1}}^{e_p} [\sin nv\,\{\cot v/2 - 2/v\} + \cos nv]/v^{\frac{1}{2}}\,dv, \end{aligned}$$

where n is the largest integer less than y, takes the place of (6.24).

[7] For the definition, properties and references concerning $\gamma_{\beta+1}$, see A, p. 416.

7.1. **Proof of Theorem** IV. Turning now to the proof of the theorem, we first select an infinite sequence of functions $\{\Phi_n(u)\}$, $n = 0, 1, \cdots$ on the basis of Lemma 1.

In the first step we set $n_0 = 0$ and choose Φ_0 so that Φ_0 has the properties specified for I_{n_0} and satisfies

$$| \Phi_0 | < 1, \qquad t_0 = \max_{0<x} | R_\alpha(x; \Phi_0) | < 1.$$

In the second step, noting that $0 < t_0$ and that

$$R_\alpha(x; \Phi_0) \to 0, \qquad x \to \infty,$$

since Φ_0 vanishes for $| u | \leqq \frac{1}{2}\pi$, we choose an integer $n_0 < n_1$ so that

$$| R_\alpha(x; \Phi_0) |/t_0 < \tfrac{1}{2}$$

for $n_1 + 1 \leqq x$. We then choose Φ_1 so that Φ_1 has the properties specified for I_{n_1} and satisfies

$$| \Phi_1 | < 1, \qquad t_1 = \max_{0<x} | R_\alpha(x; \Phi_1) | < \tfrac{1}{2} t_0.$$

Continuing this process, in the $(p + 1)$-th step we select $n_{p-1} < n_p$ so that

$$\sum_{q=0}^{p-1} | R_\alpha(x; \Phi_q) |/t_q < 2^{-p} \quad \text{for} \quad n_p + 1 \leqq x. \tag{7.11}$$

We then choose Φ_p so as to have the properties specified for I_{n_p} and satisfy

$$| \Phi_p | < 1, \qquad t_p = \max_{0<x} | R_\alpha(x; \Phi_p) | < t_{p-1}/2^p.$$

The function $f(u, v)$ we now define as

$$f(u, v) = \sum_{p=0}^{\infty} \Phi_p(u)\, \varphi_p(v),$$

where the φ's are the functions of Lemma 3 with

$$e_0 = \pi, \qquad e_1 = t_0^2, \qquad e_2 = t_1^2, \qquad \cdots.$$

Since $0 < t_{p+1} < t_p/2^{p+1} < \frac{1}{2}$, these e's satisfy (6.21). On the other hand, Φ_p, φ_p are even, periodic and integrable, and

$$\Phi_p(u) = 0, \qquad \sum_{p=0}^{\infty} | \Phi_p(u) \phi_p(v) | < v^{-\frac{1}{2}},$$

the former for $0 \leqq u \leqq \frac{1}{2}\pi$, the latter on $(0, 0; \pi, \pi)$. Hence f has all the properties asserted save perhaps those concerning summability.

We consider the Riesz mean

$$R_{\alpha,\beta}(x, y) = \sum_{m<x} (1 - m/x)^{\alpha} \sum_{n<y} (1 - n/y)^{\beta} a_{m,n}.$$

We shall show that $R_{\alpha,\beta}$ is bounded for $0 < x, 0 < y$ and that

$$\overline{\lim_{(x,y)\to(\infty,\infty)}} | R_{\alpha,\beta} | > 0. \tag{7.12}$$

From these two facts the theorem follows. For first, by Theorem II of A, boundedness $(R; \alpha, \beta)$ implies both boundedness $(C; \alpha, \beta)$ and the equivalence of summability $(R; \alpha, \beta)$ with sum s to summability $(C; \alpha, \beta)$ with sum s. And, secondly, by Theorem III, since f is continuous $(C; 0, 0)$ with limit 0, the series (5.11) cannot be summable at all unless it is summable to 0.

Now

$$R_{\alpha,\beta}(x, y) = \sum_{q=0}^{\infty} R_{\alpha}(x; \Phi_q) R_{\beta}(y; \varphi_q).$$

In addition, for $x \leqq n_{p+1} + 1, p \leqq q$,

$$R_{\alpha}(x; \Phi_q) = 0;$$

and accordingly, for $x \leqq n_{p+1} + 1$,

$$R_{\alpha,\beta} = \sum_{q=0}^{p} R_{\alpha}(x; \Phi_q) R_{\beta}(y; \varphi_q).$$

Now, for $n_1 + 1 \leqq n_p + 1 \leqq x, 0 < y$, we have

$$\left| \sum_{q=0}^{p-1} R_{\alpha}(x; \Phi_q) R_{\beta}(y; \varphi_q) \right| < M \sum_{q=0}^{p-1} | R_{\alpha}(x; \Phi_q) |/t_q < M/2^p$$

by (6.22) and (7.11). On the other hand, for $0 \leqq p, 0 < x, 0 < y$,

$$| R_{\alpha}(x; \Phi_p) R_{\beta}(y; \varphi_p) | < t_p M/t_p = M.$$

We conclude that $R_{\alpha,\beta}$ is bounded for $0 < x, 0 < y$. In addition, choosing x_p so that

$$t_p = R_{\alpha}(x_p ; \Phi_p),$$

we have

$$n_p + 1 \leqq x_p \leqq n_{p+1} + 1.$$

Hence, using (6.23) and choosing ξ so that

$$\int_{\xi}^{\infty} \gamma_{\beta+1}(v)/v^{\frac{1}{2}} dv \neq 0,$$

we have

$$\overline{\lim_{(x,y)\to(\infty,\infty)}} | R_{\alpha,\beta}(x, y) | \geqq \lim_{p\to\infty} | R_{\alpha}(x_p; \Phi_p) R_{\beta}(\xi/t_p^2; \varphi_p) |$$

$$= \left| 2 \xi^{\frac{1}{2}}/\pi \int_{\xi}^{\infty} \gamma_{\beta+1}(v)/v^{\frac{1}{2}} dv \right| \neq 0.$$

This is (7.12) and completes the proof.

8.1. **Lemmas for Theorem** V.

LEMMA 4. *Let* $0 < \alpha < a - 1$. *Let* $\varphi(u)$ *be integrable over* $(0, \pi)$, *even, and with period* 2π. *Let* (1.11) *be the Fourier series of* φ *at the origin, and let it be bounded* (C, α) *or* (R, α). (*The two are equivalent.*) *Let* m_0 *be an arbitrary positive integer. Then we can write*

$$\varphi_a(x) = \sum_{p<m_0} g_p(x)\, a_p + g_{m_0}(x),$$

where the g*'s are measurable for* $0 < x$, $g_0, \cdots, g_{m_0-1}$ *are independent of* φ, *and*

$$| g_p(x) | < M^* x^a \text{ for } p < m_0, \quad | g_{m_0}(x) | \leqq M^* x^a \operatorname*{l.u.b.}_{m_0 \leqq u} | \sigma_\alpha(u)/u^\alpha |,$$

M^* *being independent of* φ *as well as* x *and* p. *In addition, an expansion of the same type is valid with*

$$| g_{m_0}(x) | \leqq M^* x^a \operatorname*{l.u.b.}_{m_0 \leqq m} | S_\alpha(m)/m^\alpha |.$$

Denoting by ψ_a the number and $H(u)$ the function of paragraph 9.1 of A, and applying Lemma 11 of A, we have

$$\begin{aligned}
\varphi_a(x) &= \psi_a x^{\alpha+a+1} \int_0^\infty H(xu)\, \sigma_\alpha(u)\, du \\
&= \sum_{p<m_0} \left\{ \psi_a x^{\alpha+a+1} \int_p^{m_0} H(xu)(u-p)^\alpha\, du \right\} a_p + \psi_a x^{\alpha+a+1} \int_{m_0}^\infty H(xu)\sigma_\alpha(u)\, du \\
&= \sum_{p<m_0} g_p(x)\, a_p + g_{m_0}(x),
\end{aligned}$$

say. Now by Lemma 10 of A,

$$\int_p^\infty | H(xu)\,(u-p)^\alpha |\, du < x^{-\alpha-1} \int_0^\infty u^\alpha \,| H(u) |\, du < M^*/x^{\alpha+1},$$

$$\int_{m_0}^\infty | H(xu)\, \sigma_\alpha(u) |\, du \leqq x^{-\alpha-1} \operatorname*{l.u.b.}_{m_0 \leqq u} | \sigma_\alpha(u)/u^\alpha |.$$

Thus, since $H(xu)(u-p)^\alpha$, $H(xu)\sigma_\alpha(u)$ are superficially measurable, the first part of the theorem holds.

To obtain the second part we use Theorem II. We have

$$\begin{aligned}
\varphi_a(x) = \sum_{p<m_0} \left\{ \psi_a x^{\alpha+a+1} \int_p^\infty H(xu) V(u-p)\, du \right\} S_\alpha(p) \\
+ \psi_a x^{\alpha+a+1} \int_{m_0}^\infty H(xu) \sum_{m_0 \leqq p < u} V(u-p) S_\alpha(p)\, du;
\end{aligned}$$

and the second part follows from the properties of V and the definition of $S_\alpha(p)$.

8.2. LEMMA 5. *If* $0 \leqq \alpha < \xi$, $0 \leqq \beta < \eta$, *and if the series* (5.11) *is ultimately*

bounded both $(R; \alpha, \eta)$ and $(R; \xi, \beta)$, it is ultimately bounded $(R; \xi, \eta)$. In addition, the corresponding result with C in place of R holds.[8]

Consider the Riesz sum

$$\sigma_{\xi,\eta}(x, y) = \sum_{p<x} (x - p)^\alpha \sum_{q<y} (y - q)^\eta a_{p,q}.$$

Under our hypotheses there is a positive m_0 such that

$$|\sigma_{\alpha,\eta}(x, y)| < Mx^\alpha y^\eta, \qquad |\sigma_{\xi,\beta}(x, y)| < Mx^\xi y^\beta, \qquad \text{for } m_0 \leqq x,\ m_0 \leqq y.$$

Now, for $m_0 \leqq x$, $m_0 \leqq y$ we have

$$\Gamma(\alpha + 1)\Gamma(\xi - \alpha)\sigma_{\xi,\eta}/\Gamma(\xi + 1) = \int_0^x (x - t)^{\xi-\alpha-1} \sigma_{\alpha,\eta}(t, y)\, dt$$

$$(8.21) \qquad = \sum_{p<m_0} \int_p^{m_0} (x - t)^{\xi-\alpha-1} (t - p)^\alpha\, dt \sum_{q<y} (y - q)^\eta a_{p,q} + \int_{m_0}^x (x - t)^{\xi-\alpha-1} \sigma_{\alpha,\eta}(t, y)\, dt$$

$$= \sigma^{(1)} + \sigma^{(2)},$$

say, and

$$(8.22) \qquad |\sigma^{(2)}| < M\, x^\xi\, y^\eta,$$

$$|\sigma^{(1)}| < M\, |\sigma_{\xi,\eta}| + |\sigma^{(2)}|$$

$$(8.23) \qquad < M \sum_{q<m_0} \int_q^{m_0} (y - q)^{\eta-\beta-1} (t - q)^\beta\, dt \sum_{p<x} (x - p)^\xi\, |a_{p,q}| + Mx^\xi y^\eta$$

$$< F(x)\, y^\eta,$$

where F is independent of y. But

$$\sigma^{(1)}\, x^{-\xi} y^{-\eta} = \sum_{p<m_0} \left\{ x^{-\xi} \int_p^{m_0} (x - t)^{\xi-\alpha-1} (t - p)^\alpha\, dt \right\} \left\{ \sum_{q<y} (1 - q/y)^\eta a_{p,q} \right\}$$

is a sum of the form considered by Agnew in his lemma on double series,[9] and

$$x^{-\xi} \int_p^{m_0} (x - t)^{\xi-\alpha-1} (t - p)^\alpha\, dt = o(1).$$

We conclude from the lemma and (8.23) that $\sigma^{(1)}$ satisfies (8.22). The proof is then complete for Riesz sums.

The proof for Cesàro sums is analogous, the formula taking the place of (8.21) being

[8] The fact that (5.11) is a Fourier series is of no importance.

[9] For a statement of and references to Agnew's lemma see A, pp. 402 and 409. As stated in A, Agnew considered the case in which x, y are integral variables. The immediate extension to continuous variables is given on p. 409 of A.

$$S_{\xi,\eta}(m, n) = \sum_{p=0}^{m} A_{n-p}^{\xi-\alpha-1} S_{\alpha,\eta}(p, n).$$

9.1. **Proof of Theorem** V. Since continuity $(C; a, b)$ implies continuity $(C; a', b')$ for $a \leqq a'$, $b \leqq b'$, we can suppose that

$$0 \leqq \alpha < a - 1 < \xi - 1 \leqq h, \qquad 0 \leqq \beta < b - 1 < \eta - 1 \leqq k,$$

where h is the integral part of a, k the integral part of b. Also, for the usual reasons, we can assume that s is 0.

Consider then $\sigma_{\xi,\eta}$. Let $0 < \epsilon$ be arbitrary. Let $0 < \delta$ be chosen so that

$$|f_{a,b}(x, y)| < \epsilon x^a y^b \qquad \text{for } x \leqq \delta, y \leqq \delta.$$

Let ψ, ψ_a, ψ_b be the numbers and $H(u)$, $K(v)$ the functions of paragraph 7.2 of A with ξ in place of α and η in place of β. Then we have, by Lemma 9 of A,

$$\begin{aligned}\sigma_{\xi,\eta} &= \psi\, x^{\xi+a+1} y^{\eta+b+1} \int_{(0,0;\infty,\infty)} H(xu)\, K(yv)\, f_{a,b}(u, v)\, d(u, v)\\ &= \left[\int_{(0,0;\delta,\delta)} - \int_{(\delta,\delta;\infty,\infty)} + \int_{(\delta,0;\infty,\infty)} + \int_{(0,\delta;\infty,\infty)}\right]\\ &\qquad\qquad \psi x^{\xi+a+1} y^{\eta+b+1} HK f_{a,b}\, d(u, v)\\ &= \sum_{i=1}^{4} r_i,\end{aligned}$$

say. Now by Lemma 6 of A,

$$(9.11) \qquad |r_1| < Nx^\xi y^\eta \epsilon \int_0^\infty |H(u)|\, u^a\, du \int_0^\infty |K(v)|\, v^b\, dv < Nx^\xi y^\eta \epsilon,$$

where N is independent of ϵ, x, y. Next, since $1 < a$, $1 < b$, we have

$$|f_{a,b}(x, y)| < Mx^a y^b \qquad \text{for } \delta \leqq x, \delta \leqq y,$$

and thus

$$(9.12) \qquad |r_2| < Mx^a y^b \int_\delta^\infty u^{a-\xi-1}\, du \int_\delta^\infty v^{b-\eta-1}\, dv < Mx^a y^b.$$

It is enough then to show in the case of R summability that $r_3 = o(x^\xi y^\eta)$, $r_4 = o(x^\xi y^\eta)$. Moreover, since the situation is symmetrical we can confine ourselves to r_3.

We first apply Lemma 5, choosing a positive integer m_0 so that

$$|\sigma_{\xi,\eta}(x, y)| < Mx^\xi y^\eta \qquad \text{for } m_0 \leqq x, n_0 \leqq y.$$

Then we note that

$$
\begin{aligned}
| r_3 | &\leqq \{| \sigma_{\xi, \eta} | + | r_1 | + | r_2 |\} + | r_4 | \\
&= O\left[x^{\xi} y^{\eta} + \left| y^{\eta+b+1} \int_{\delta}^{\infty} K(yv) \left\{ \sum_{p<x} (x - p)^{\xi} \lambda_p \right.\right.\right. \\
&\qquad \left.\left.\left. \int_0^{\pi} \cos pu\, f_{0,b}(u, v)\, du \right\} dv \right| \right] \qquad (9.13) \\
&= O\left\{x^{\xi} y^{\eta} + x^{\xi+1} y^{\eta+b+1} \int_{\delta}^{\infty} | K(yv) |\, dv \int_0^{\pi} | f_{0,b}(u, v) |\, du\right\} \\
&= O\,(x^{\xi+1} y^{\eta}).
\end{aligned}
$$

Next, we write

$$
r_3 = \psi_a x^{\xi+a+1} \int_{\delta}^{\infty} H(xu) \left\{ \sum_{q<y} (y - q)^{\eta} \lambda_q \int_0^{\pi} \cos qv\, f_{a,0}(u, v)\, dv \right\} du.
$$

Setting

$$
\varphi(u, y) = \sum_{q<y} (y - q)^{\eta} \lambda_q \int_0^{\pi} \cos qv\, f(u, v)\, dv,
$$

we have, for a fixed y,

$$
\varphi_a(u, y) = \sum_{q<y} (y - q)^{\eta} \lambda_q \int_0^{\pi} \cos qv\, f_{a,0}(u, v)\, dv.
$$

In addition, φ is an even periodic function of u. Its Fourier series at the origin is

$$
\sum_{p=0}^{\infty} \lambda_p \left\{ \sum_{q<y} (y - q)^{\eta} \lambda_q a_{p,q} \right\},
$$

which, for $m_0 \leqq y$, is bounded (R, α). Accordingly, using Lemma 4, we can write

$$
\varphi_a(u, y) = \sum_{p<m_0} g_p(u) \sum_{q<y} (y - q)^{\eta} \lambda_q a_{p,q} + g(u, y) \qquad \text{for } m_0 \leqq y,
$$

where the g's are measurable for $0 < u$, $g_0, \cdots, g_{m_0-1}$ are independent of y, and

$$
| g_p(x) | < Mx^{a} \text{ for } p < m_0, \qquad | g(x, y) | < Mx^{a} y^{\eta} \text{ for } m_0 \leqq y.
$$

Thus,

$$
\begin{aligned}
r_3 &- \sum_{p<m_0} \psi_a x^{\xi+a+1} \int_{\delta}^{\infty} H(xu)\, g_p(u)\, du \sum_{q<y} (y - q)^{\eta} \lambda_q a_{p,q} \\
&= O\left\{ x^{\xi+a+1} \left| \int_{\delta}^{\infty} g(u, y)\, H(xu)\, du \right| \right\} \\
&= O\,(x^{a} y^{\eta}) = o\,(x^{\xi} y^{\eta}).
\end{aligned}
$$

But the sum

$$(9.14)\quad \sum_{p<m_0}\left\{\psi_a x^{a+1}\int_\delta^\infty H(xu)g_p(u)\,du\right\}\left\{\sum_{q<y}(1-q/y)^\eta\lambda_q a_{p,q}\right\} = r_3\,x^{-\xi}y^{-\eta} + o(1)$$

is again one to which Agnew's lemma is applicable. Since

$$\psi_a x^{a+1}\int_\delta^\infty H(xu)g_p(u)\,du = O(x^{a-\xi}) = o(1),$$

we conclude, on applying (9.13) and the lemma, that the sum is $o(1)$. From (9.14) we then see that $r_3 = o(x^\xi y^\eta)$, and this completes the proof for Riesz sums.

The proof for Cesàro sums is of the same type but slightly more complicated. We first choose $f, \bar{f}$ so as to satisfy the conditions imposed upon f in Theorem I with α replaced by ξ, η. Then we have

$$S_{\xi,\eta}(m,n) = \int_0^{m+1} U(m+1-x)dx\int_0^{n+1}\bar{U}(n+1-y)\,\sigma_{\xi,\eta}(x,y)\,dy,$$

where U, $\bar{U}$ are the fractional derivatives $f^{(\xi+1)}$, $\bar{f}^{(\eta+1)}$. Letting ϵ, δ, r_i have the same significance as above, we write

$$S_{\xi,\eta} = \sum_{i=1}^{4}\int_0^{m+1} U(m+1-x)dx\int_0^{n+1}\bar{U}(n+1-y)\,r_i\,dy = \sum_{i=1}^{4} S_i,$$

say.

We see readily by (9.11) that

$$|S_1| < N(m+1)^\xi(n+1)^\eta\epsilon\int_0^{m+1}|U(x)|\,dx\int_0^{n+1}|U(y)|\,dy$$
$$< N(m+1)^\xi(n+1)^\eta\epsilon,$$

where N is independent of m, n, ϵ. In addition, we have, by (9.12),

$$|S_2| < M(m+1)^a(n+1)^b = o(m^\xi n^\eta).$$

Accordingly, because of the symmetry, it is enough to show that $S_3 = o(m^\xi n^\eta)$.

We first apply Lemma 5, choosing a positive integer m_0 so that

$$|S_{\xi,\eta}| < Mm^\xi n^\eta \qquad \text{for } m_0 \leqq m,\ m_0 \leqq n.$$

Then we note that

$$|r_4| \leqq M(x+1)^{\xi+1}y^{\eta+b+1}\int_\delta^\infty|K(yv)|\,dv\int_0^\pi|f_{0,b}(u,v)|\,du$$
$$< M(x+1)^{\xi+1}y^\eta,$$

and, accordingly, that

$$
\begin{aligned}
S_3 = O\Big\{ | S_{\xi,\eta} | + | S_1 | + S_2 | &+ \int_0^{m+1} | U(m+1-x) | \, dx \\
&\int_0^{n+1} | \overline{U}(n+1-y) \, r_4 | \, dy \Big\}
\end{aligned} \tag{9.15}
$$

$$
= O\{m^{\xi+1} n^{\eta}\}.
$$

Next, we write

$$
\int_0^{n+1} \overline{U}(n+1-y) \, r_3 \, dy = \psi_a x^{\xi+a+1} \int_\delta^\infty H(xu) \left\{ \sum_{q=0}^{n} A^{\eta}_{n-q} \lambda_q \int_0^\pi \cos qv \, f_{a,0}(u, v) \, dv \right\} du;
$$

and reasoning as with r_3 we see that

$$
\begin{aligned}
\int_0^{n+1} \overline{U} r_3 \, dy - \sum_{p<m_0} \psi_a x^{\xi+a+1} \int_\delta^\infty H(xu) g_p^*(u) \, du \sum_{q=0}^{n} A^{\eta}_{n-q} \lambda_q \, a_{p,q} \\
= \psi_a x^{\xi+a+1} \int_\delta^\infty H(xu) \, g^*(u, n) \, du,
\end{aligned}
$$

where the g^*'s have properties analogous to the g's. Thus, for $m_0 \leqq n$,

$$
\begin{aligned}
S_3 - \sum_{p<m_0} \left\{ \psi_a \int_0^{m+1} U(m+1-x) x^{\xi+a+1} \, dx \int_\delta^\infty H(xu) g_p^*(u) \, du \right\} \left\{ \sum_{q=0}^{n} A^{\eta}_{n-q} \lambda_q \, a_{p,q} \right\} \\
= \psi_a \int_0^{m+1} U(m+1-x) x^{\xi+a+1} \, dx \int_\delta^\infty H(xu) g^*(u, n) \, du \\
= O\left\{ n^{\eta} \int_0^{m+1} | U | \, x^a \, dx \int_\delta^\infty u^{a-\xi-1} \, du \right\} O(m^a n^{\eta}).
\end{aligned}
$$

Making use of (9.15) and applying Agnew's lemma to the sum

$$
\sum_{p<m_0} \left\{ m^{-\xi} \int_0^{m+1} U \, x^{\xi+a+1} \, dx \int^\infty H g_p^* \, du \right\} \left\{ \sum_{q=0}^{n} A^{\eta}_{n-q} n^{-\eta} \lambda_q \, a_{p,q} \right\},
$$

we conclude that $S_3 = o(m^{\xi} n^{\eta})$, and this completes the proof.

UNIVERSITY OF ROCHESTER AND DUKE UNIVERSITY.

STRUCTURES AND GROUP THEORY. I

By Oystein Ore

In a recent paper on the foundations of abstract algebra[1] I have shown that the principal results on algebraic domains are not primarily to be considered as properties of the elements of the domain itself but as properties of certain systems of distinguished subsets, like systems of subgroups, ideals, submoduli, etc. These systems of subsets have the common characteristic property that they form a *structure*, i.e., a system in which *union* and *cross-cut* of two elements are defined. The theorems on algebraic domains are shown to be theorems on structures. This explains the well-known similarity of several algebraic theories and makes possible a unified structural theory applicable to all systems.

After this common foundation for the algebraic theories has been established, it is, however, not difficult to see that the various algebraic domains like fields, rings and groups have peculiar structural properties of their own. In certain cases it is even possible to characterize the domains by these properties.

In the following we shall apply the principles of the theory of structures to the foundation of the theory of groups, that is, we base the theory as far as possible directly upon the properties of subgroups and eliminate the elements from theorems and proofs. This entails a certain simplification. More important, however, is the fact that this method, even in the elementary theory of groups which we consider in this paper, leads to new systematic points of view and interesting new results.

In Chapter I we discuss the structure formed by all subgroups of a given group and indicate the general principle of duality. Furthermore, in the theory of structures we have constructed a quotient structure for any structure, while quotient systems A/B in groups have been defined only when B is normal in A. Hence we are led to the introduction of quotient systems for all subgroups. The algebraic system A/B is then a *multi-group* differing from ordinary groups only in the property that the product is not unique.

In Chapter II we consider the law of isomorphism. When the assumption of isomorphism is weakened to co-set correspondence we are led to permutable groups. Such groups have the structural property expressed in Theorem 6. When structure isomorphism is required, one is led to the important type of subgroups which I have called quasi-normal subgroups.

Some of the principal new results are to be found in Chapter III. The

Received February 10, 1937.

[1] Oystein Ore, *On the foundation of abstract algebra* I, Annals of Math., vol. 36 (1935), pp. 406–437; II, ibid., vol. 37 (1936), pp. 265–292. These two papers will be cited as *Foundations* I and II.

theorems of Jordan-Hölder and Schreier-Zassenhaus are analysed and several extensions of these theorems are obtained. When only index relations are required certain weak permutability conditions of the chains are required. When structure isomorphism is wanted we have to consider chains in which every term is quasi-normal in the preceding. For such chains there exists a complete analogue of the theorem of Jordan-Hölder.

The last chapter deals with the Dedekind structure formed by all normal subgroups. Here some of the results may be taken over directly from the theory of structures. The concept of direct similarity is shown to be equivalent to central isomorphism. The well-known duality between the center and the anti-center, i.e., the quotient group with respect to the commutator group is easily explained by the general principle of duality. Among the further results of this chapter I shall only mention the interesting self-dual Theorem 4.

Chapter 1. Quotient systems

1. **Group structures.**[2] The system of all subgroups of a given group G forms a *structure* Σ. To any two given subgroups A and B of G there exists a *cross-cut* (A, B) consisting of the common elements of A and B, and a *union* $[A, B]$, which is the subgroup generated by A and B. Since this structure has the property that any finite or infinite set of subgroups has a cross-cut and union, we shall say *the structure Σ of all subgroups of a given group G is closed.*

The cross-cut and union satisfy the ordinary axioms for these operations:

$$(A, B) = (B, A), \qquad [A, B] = [B, A],$$
$$(A, A) = A, \qquad [A, A] = A,$$
$$(A, (B, C)) = ((A, B), C), \qquad [A, [B, C]] = [[A, B], C],$$
$$(A, [B, A]) = A, \qquad [A, (B, A)] = A.$$

It is obvious that the two operations correspond dualistically. This simple remark enables us to express an important principle which we shall repeatedly apply.

Principle of Duality. *To any structure theorem on groups corresponds a dual obtained by interchanging the two concepts union and cross-cut.*

In the following we shall consider various substructures of the structure Σ of all subgroups. Two structures Σ_1 and Σ_2 shall be said to be *structure isomorphic*, $\Sigma_1 \sim \Sigma_2$ if there exists a one-to-one correspondence between the subgroups of the two structures such that if

$$A_1 \rightleftarrows A_2, \qquad B_1 \rightleftarrows B_2,$$

then

$$(A_1, B_1) \rightleftarrows (A_2, B_2), \qquad [A_1, B_1] \rightleftarrows [A_2, B_2].$$

If $A > B$ are two subgroups of G, we shall denote the *index* of B in A by

[2] Compare *Foundation* I, Chapter 1.

$\{A:B\}$. The two structures Σ_1 and Σ_2 shall be said to be *strongly structure isomorphic* if for any two subgroups $A_1 > B_1$ and the corresponding $A_2 > B_2$ we also have the same index

$$\{A_1:B_1\} = \{A_2:B_2\}.$$

In this case we shall write $\Sigma_1 \simeq \Sigma_2$.

2. **Quotient systems.**[3] In an arbitrary structure Σ any two elements $A \geqq B$ define a substructure consisting of all elements H between A and B

$$A \geqq H \geqq B.$$

This structure shall be called the *quotient structure* of A and B and we shall denote it by $\mathfrak{A} = A/B$.

In the case where B is a normal subgroup of A, we ordinarily associate a *quotient group* with the symbol A/B. We shall now show that in the non-normal case it is also possible to associate with the quotient A/B an algebraic system which in the case of a normal subgroup reduces to the quotient group.

We represent the group A by means of its left-hand co-sets with respect to B:

$$A = \{a_1B, a_2B, \cdots\} = a_1B + a_2B + \cdots, \tag{1}$$

where the a_i are suitably chosen elements of A. It should be observed that it is sufficient to consider only one set of co-sets (1), since the right-hand co-set representation may be obtained from (1) by the inverse automorphism $a \to a^{-1}$ in A:

$$A = \{Ba_1^{-1}, Ba_2^{-1}, \cdots\} = Ba_1^{-1} + Ba_2^{-1} + \cdots.$$

The totality of left-hand co-sets (1) shall be called the *quotient system* of A with respect to B and we shall write

$$A/B = \{B_1, B_2, \cdots\}, \tag{2}$$

where the $B_i = a_iB$ denote the co-sets in some order. This quotient set will now be made into an algebraic system through the definition of a *multiplication*. This multiplication differs, however, from the ordinary multiplication in algebraic systems by the property that two elements do not define a unique product element, but a *product set* consisting of several elements. Let B_α and B_β be two co-sets. The elements of A contained in the complex

$$B_\alpha \cdot B_\beta = a_\alpha \cdot B \cdot a_\beta \cdot B \tag{3}$$

will then give a certain subset of co-sets in (2)

$$S = \{B_{\gamma_1}, B_{\gamma_2}, \cdots\},$$

where each B_γ is of the form

[3] *Foundation* I, Chapter 3.

$$B_\gamma = a_\alpha \cdot b \cdot a_\beta \cdot B,$$

and b is an element of B. We now define the product of two co-sets as

$$B_\alpha \cdot B_\beta = \{B_{\gamma_1}, B_{\gamma_2}, \cdots\}. \tag{4}$$

It is natural to extend this definition of the product of co-sets to the product of arbitrary subsets

$$S = \{B_{\gamma_1}, \cdots\}, \qquad T = \{B_{\delta_1}, \cdots\}$$

of A/B by saying that $S \cdot T$ is the set of co-sets containing all products $B_\gamma \cdot B_\delta$. This multiplication is obviously associative.

Let us now mention a few of the properties of the new multiplication. The product $B_\alpha \cdot B_\beta$ contains the co-set $a_\alpha \cdot a_\beta \cdot B$. In the multiplication, the group B plays the rôle of a *right-hand unit element* with the properties

$$B_\alpha \cdot B = B_\alpha, \qquad B \cdot B_\alpha > B_\alpha$$

for all co-sets B_α. One may define an *inverse* of a co-set B_α as a co-set B_β such that

$$B_\beta \cdot B_\alpha > B. \tag{5}$$

This relation is equivalent to the existence of elements b_1 and b_2 in B such that

$$a_\beta \cdot b_1 \cdot a_\alpha = b_2. \tag{6}$$

This relation may always be satisfied by taking b_1 equal to the unit element e and $a_\beta = a_\alpha^{-1}$. By taking the inverse of the relation (6), one finds that (5) also implies

$$B_\alpha \cdot B_\beta > B.$$

Hence we have shown

THEOREM 1. *Every element of the quotient system A/B has at least one inverse, and if B_β is an inverse of B_α, then B_α is an inverse of B_β.*

Let us determine when a product is unique. If $a_\alpha \cdot B$ and $a_\beta \cdot B$ have a unique product, we must have

$$a_\alpha \cdot B \cdot a_\beta \cdot B = a_\alpha \cdot a_\beta \cdot B,$$

and hence for every b_1 in B there must exist another b_2 such that

$$a_\alpha \cdot b_1 \cdot a_\beta = a_\alpha \cdot a_\beta \cdot b_2$$

or

$$a_\beta^{-1} \cdot b_1 \cdot a_\beta = b_2.$$

This shows

THEOREM 2. *A product $B_\alpha \cdot B_\beta$ is unique if and only if B_β belongs to the normaliser group of B in A.*

Similarly one proves

THEOREM 3. *The normaliser group of B with respect to A consists of all co-sets with a unique inverse.*

3. **Multigroups.** The properties of a quotient system naturally lead to the definition of a new abstract system which may be called a *multigroup*. It may be defined as a system $\mathfrak{M}$ of elements satisfying the following axioms:

I. The product of two elements B_α and B_β is a subset of $\mathfrak{M}$:

$$B_\alpha \cdot B_\beta = \{B_{\gamma_1}, B_{\gamma_2}, \cdots\}.$$

More generally, the product of two sets of elements

$$S = \{B_{\alpha_1}, B_{\alpha_2}, \cdots\}, \qquad T = \{B_{\beta_1}, B_{\beta_2}, \cdots\}$$

is the set defined by all products $B_\alpha \cdot B_\beta$.

II. The multiplication is associative.

III. There exists a unique right-hand unit element B_1 such that for all B_α

$$B_\alpha \cdot B_1 = B_\alpha, \qquad B_1 \cdot B_\alpha > B_\alpha.$$

IV. To each B_α there exists at least one inverse B'_α such that

$$B_\alpha \cdot B'_\alpha > B_1, \qquad B'_\alpha \cdot B_\alpha > B_1.$$

This definition shows that a multigroup differs from an ordinary group only in the property that the product is not unique. Obviously any quotient system A/B is a multigroup. In the case of a multigroup arising from co-sets in a group there are further axiomatic conditions satisfied which have not been enumerated above. Although these multigroups have several interesting properties, the study of these must be reserved for a later occasion. I shall only mention here some recent investigations by Marty[4] on the subject. The multigroups are closely related to the "Mischgruppen" studied by Loewy.[5] It seems to me that the representation of the quotient systems by means of multigroups is preferable to the representation by means of "Mischgruppen", since one avoids the double system of operators which the latter theory introduces.

We shall complete these remarks by proving one theorem which is of importance for the sequel. Let $A > B$ be two subgroups of G and A_1 any subgroup of A containing B. The group A_1 consists of certain co-sets of A/B and these co-sets are seen to form a submultigroup of A/B. We shall show conversely that every submultigroup A_1 of A/B constitutes a subgroup of A containing B when considered as a set of elements in A. According to the definition

[4] F. Marty, *Sur les groupes et hypergroupes attachès à une fonction rationelle*, Annales de l'Ecole Normale Supérieure, (3), vol. 53 (1936), pp. 83–123. A list of further publications on this subject by the same author is given in the introduction. After the completion of this manuscript appeared a paper by H. S. Wall, *Hypergroups*, Am. Journal of Math., vol. 59 (1937), pp. 77–98.

[5] A. Loewy, *Über abstrakt definierte Transmutationssysteme oder Mischgruppen*, Journal für Mathematik, vol. 157 (1927), pp. 239–254. See also R. Baer, Sitzungsberichte, Heidelberg, 1928.

A_1 contains the product of any of its co-sets. Hence according to a preceding remark it contains the product of any of its elements. To prove that A_1 contains the inverse a_1^{-1} of any of its elements a_1, we recall that A_1 contains an inverse co-set $a_2 \cdot B$ to $a_1 \cdot B$. As in (6), we then find $a_2 \cdot b_1 \cdot a_1 = b_2$, where b_1 and b_2 belong to B. This shows that a_1^{-1} belongs to the co-sets $B \cdot a_2 \cdot B$ which are contained in A_1. Hence we have

THEOREM 4. *The submultigroups of a quotient system A/B when considered as a set of elements in A are identical with the subgroups of A containing B.*

4. **Concepts in group structures.**[6] We shall now define various concepts which are of importance for the theory of group structures. We consider as before the structure Σ of all subgroups of a given group G. A *unit quotient* in Σ is any quotient

$$\mathfrak{E} = A/A.$$

The *product of quotients* may be introduced in the following manner: If we have

$$\mathfrak{A} = A/B, \qquad \mathfrak{B} = B/C, \qquad \mathfrak{C} = A/C,$$

we shall write

$$\mathfrak{C} = \mathfrak{A} \times \mathfrak{B}, \qquad \mathfrak{A}^{-1} \times \mathfrak{C} = \mathfrak{B}, \qquad \mathfrak{C} \times \mathfrak{B}^{-1} = \mathfrak{A},$$

and say that $\mathfrak{A}$ is a *left-hand* and $\mathfrak{B}$ a *right-hand factor* of $\mathfrak{C}$. It is often convenient to consider any subgroup A or even the group G itself as a quotient with respect to the unit element

$$\mathfrak{A} = A/E, \qquad \mathfrak{G} = G/E.$$

The system of all quotients may be made into a structure, the *quotient structure of* Σ. When

$$\mathfrak{A}_1 = A_1/B_1, \qquad \mathfrak{A}_2 = A_2/B_2,$$

we define

$$(\mathfrak{A}_1, \mathfrak{A}_2) = (A_1, A_2)/(B_1, B_2), \qquad [\mathfrak{A}_1, \mathfrak{A}_2] = [A_1, A_2]/[B_1, B_2].$$

We shall apply these concepts mainly in the case where $\mathfrak{A}_1$ and $\mathfrak{A}_2$ have the same denominator

$$\mathfrak{A}_1 = A_1/B, \qquad \mathfrak{A}_2 = A_2/B, \tag{7}$$

and hence

$$(\mathfrak{A}_1, \mathfrak{A}_2) = (A_1, A_2)/B, \qquad [\mathfrak{A}_1, \mathfrak{A}_2] = [A_1, A_2]/B. \tag{8}$$

The two quotients $\mathfrak{A}_1$ and $\mathfrak{A}_2$ will be said to be relatively prime when their cross-cut is a unit quotient, i.e., when

$$(A_1, A_2) = B.$$

[6] *Foundations* I, Chapter 3.

We may also apply the principle of duality and define the duals of these concepts. Corresponding to (7) we consider quotients with the same numerator

(9) $$\mathfrak{B}_1 = A/B_1, \qquad \mathfrak{B}_2 = A/B_2.$$

We define their *left-hand union and cross-cut* respectively as

$$[\mathfrak{B}_1, \mathfrak{B}_2]_l = A/(B_1, B_2), \qquad (\mathfrak{B}_1, \mathfrak{B}_2)_l = A/[B_1, B_2].$$

The two quotients (5) are l.h. relatively prime if their l.h. cross-cut is a unit quotient, i.e., if

$$[B_1, B_2] = A.$$

Next it is convenient to introduce the notion of *transformation*. Let

(10) $$\mathfrak{A} = A/B, \qquad \mathfrak{C} = C/B$$

be two quotients with the same denominator. The quotient

(11) $$\mathfrak{A}' = [\mathfrak{A}, \mathfrak{C}] \times \mathfrak{C}^{-1} = \mathfrak{C}\mathfrak{A}\mathfrak{C}^{-1} = [A, C]/C$$

is then said to have been obtained from $\mathfrak{A}$ through (*right-hand*) *transformation with* $\mathfrak{C}$.

The transformation has a series of simple properties which we shall now enumerate. The proofs follow directly from the definition (11).

Lemma 1. $$[\mathfrak{A}, \mathfrak{C}] = \mathfrak{C}\mathfrak{A}\mathfrak{C}^{-1} \times \mathfrak{C}.$$

Lemma 2. $$(\mathfrak{B} \times \mathfrak{C})\mathfrak{A}(\mathfrak{B} \times \mathfrak{C})^{-1} = \mathfrak{B}(\mathfrak{C}\mathfrak{A}\mathfrak{C}^{-1})\mathfrak{B}^{-1}.$$

Lemma 3. $$\mathfrak{C}[\mathfrak{A}, \mathfrak{B}]\mathfrak{C}^{-1} = [\mathfrak{C}\mathfrak{A}\mathfrak{C}^{-1}, \mathfrak{C}\mathfrak{B}\mathfrak{C}^{-1}].$$

Lemma 4. $$(\mathfrak{C} \times \mathfrak{B})(\mathfrak{A} \times \mathfrak{B})(\mathfrak{C} \times \mathfrak{B})^{-1} = \mathfrak{C}\mathfrak{A}\mathfrak{C}^{-1}.$$

Lemma 5. $$\mathfrak{C}(\mathfrak{A} \times \mathfrak{B})\mathfrak{C}^{-1} = \mathfrak{C}_1\mathfrak{A}\mathfrak{C}_1^{-1} \times \mathfrak{C}\mathfrak{B}\mathfrak{C}^{-1},$$

where

$$\mathfrak{C}_1 = \mathfrak{B}\mathfrak{C}\mathfrak{B}^{-1}.$$

This rule for the transformation of a product may be extended to an arbitrary number of factors.

Lemma 6. $$[\mathfrak{A}_1 \times \mathfrak{A}, \mathfrak{B}_1 \times \mathfrak{B}]/[\mathfrak{A}, \mathfrak{B}] = [\mathfrak{A}_2\mathfrak{A}_1\mathfrak{A}_2^{-1}, \mathfrak{B}_2\mathfrak{B}_1\mathfrak{B}_2^{-1}],$$

where

$$\mathfrak{A}_2 = \mathfrak{A}\mathfrak{B}\mathfrak{A}^{-1}, \qquad \mathfrak{B}_2 = \mathfrak{B}\mathfrak{A}\mathfrak{B}^{-1}.$$

The following theorem is of considerable importance.

Theorem 5. *If a product* $\mathfrak{B} \times \mathfrak{C}$ *has a r.h. factor* $\mathfrak{A}$, *then* $\mathfrak{B}$ *has the r.h. factor* $\mathfrak{C}\mathfrak{A}\mathfrak{C}^{-1}$.

Proof. Since $\mathfrak{B} \times \mathfrak{C}$ has the two factors $\mathfrak{A}$ and $\mathfrak{C}$, we can write

$$\mathfrak{B} \times \mathfrak{C} = \mathfrak{B}_1 \times [\mathfrak{A}, \mathfrak{C}],$$

and division by $\mathfrak{C}$ gives

$$\mathfrak{B} = \mathfrak{B}_1 \times \mathfrak{C}\mathfrak{A}\mathfrak{C}^{-1}.$$

Dually corresponding to right-hand transformation we also have *left-hand transformation*, which may be defined as follows. Let

$$(12) \qquad \mathfrak{A} = A/B, \qquad \mathfrak{D} = A/D$$

be two quotients with the same numerator. The quotient

$$(13) \qquad \mathfrak{A}'' = \mathfrak{D}^{-1} \times [\mathfrak{A}, \mathfrak{D}]_l = \mathfrak{D}^{-1}\mathfrak{A}\mathfrak{D} = D/(B, D)$$

is then the left-hand transform of $\mathfrak{A}$ by $\mathfrak{D}$.

In the special case where the quotients $\mathfrak{A}$ and $\mathfrak{C}$ in (10) are relatively prime, i.e., $[A, C] = B$, we shall say that $\mathfrak{A}'$ in (11) has been obtained from $\mathfrak{A}$ by a *r.h. similarity transformation*. In the same manner $\mathfrak{A}''$ in (13) is obtained from $\mathfrak{A}$ in (12) by a *l.h. similarity transformation* if $\mathfrak{A}$ and $\mathfrak{D}$ are l.h. relatively prime. The l.h. and r.h. similarity transformations are inverse in the sense that the l.h. similarity transformation of $\mathfrak{A}'$ in (11) by $[A, C]/A$ gives $\mathfrak{A}$, while the r.h. similarity transformation of $\mathfrak{A}''$ in (13) by $B/(B, D)$ also gives $\mathfrak{A}$.

Let us finally mention

THEOREM 6. *Any right-hand union of two r.h. relatively prime quotients with the same denominator*

$$\mathfrak{M} = [\mathfrak{A}, \mathfrak{B}], \qquad (\mathfrak{A}, \mathfrak{B}) = \mathfrak{E}$$

is also the l.h. union

$$\mathfrak{M} = [\mathfrak{A}', \mathfrak{B}']_l, \qquad (\mathfrak{A}', \mathfrak{B}')_l = \mathfrak{E}'$$

of two l.h. relatively prime quotients

$$\mathfrak{A}' = \mathfrak{B}\mathfrak{A}\mathfrak{B}^{-1}, \qquad \mathfrak{B}' = \mathfrak{A}\mathfrak{B}\mathfrak{A}^{-1}$$

obtained from $\mathfrak{A}$ and $\mathfrak{B}$ by similarity transformation.

The proof is immediate. The dual of Theorem 6 is obviously also true. Furthermore it may be extended to the union of an arbitrary number of quotients each relatively prime to the union of the rest.

Chapter 2. The law of isomorphism

1. **Correspondences.** A fundamental theorem in group theory is the ordinary *law of isomorphism*:

Let A and B be two subgroups such that A is normal in $[A, B]$. Then (A, B) is normal in B and there exists an (element) isomorphism between the two quotient groups

$$[A, B]/A \cong B/(A, B).$$

We shall now consider what remains of this law when we drop the condition of normality and consider two arbitrary subgroups A and B of G. We denote their union and cross-cut by

$$M = [A, B], \qquad D = (A, B),$$

and we wish to compare the two quotient systems

$$\mathfrak{A} = M/A, \qquad \mathfrak{B} = B/D.$$

The quotient system $\mathfrak{B}$ is made up by certain co-sets

$$\mathfrak{B} = \{\cdots, b_i D, \cdots\}. \tag{1}$$

Those co-sets of $\mathfrak{A}$ which have the same multipliers as in (1) we shall call the *co-sets of $\mathfrak{A}$ corresponding to $\mathfrak{B}$*. They form a subset

$$\mathfrak{A}' = \{\cdots, b_i A, \cdots\} \tag{2}$$

of $\mathfrak{A}$. These co-sets are all different, and hence we have the well-known

THEOREM 1. *For any two groups A and B the index relation*

$$\{[A, B]:A\} \geqq \{B:(A, B)\}$$

holds.

Obviously the corresponding co-sets (2) form a generating system for the quotient system $\mathfrak{A}$ in the sense that one obtains $\mathfrak{A}$ by taking all finite products of them. Let us now determine when they actually constitute the whole group $[A, B]$. In this case each element of $[A, B]$ must belong to some co-set $b \cdot A$ and hence each product $a_1 \cdot b_1$ also has a representation $b \cdot a$. We shall say that two groups A and B are *permutable* if

$$[A, B] = A \cdot B = B \cdot A,$$

i.e., if to any element a_1 in A and b_1 in B there exist other elements such that $a_1 \cdot b_1 = b_2 \cdot a_2$. If A and B are permutable, there exists a one-to-one correspondence $b \cdot A \rightleftarrows b \cdot D$ between $\mathfrak{A}$ and $\mathfrak{B}$. We shall call this the *regular co-set correspondence* and write $\mathfrak{A} \rightleftarrows \mathfrak{B}$.

THEOREM 2. *The necessary and sufficient condition that there exist a regular correspondence between the quotients*

$$\mathfrak{A} = [A, B]/A, \qquad \mathfrak{B} = B/(A, B)$$

is that A and B be permutable.

A consequence is

THEOREM 3. *Let the index $\{B:(A, B)\}$ be finite. The necessary and sufficient condition that*

$$\{[A, B]:A\} = \{B:(A, B)\} \tag{4}$$

is that A and B be permutable.

One may now ask when there exists a *structure isomorphism* between the two quotient systems $\mathfrak{A}$ and $\mathfrak{B}$. One can always obtain a correspondence between the subgroups of $\mathfrak{A}$ and $\mathfrak{B}$ in the following manner: Let us denote by $\bar{A}$ and $\bar{B}$ arbitrary subgroups in $\mathfrak{A}$ and $\mathfrak{B}$

$$(5)\qquad [A, B] \geqq \bar{A} \geqq A, \qquad B \geqq \bar{B} \geqq (A, B).$$

The correspondence

$$(6)\qquad \bar{A} \to (\bar{A}, B), \qquad \bar{B} \to [A, \bar{B}]$$

will then be called the *regular structure correspondence* between $\mathfrak{A}$ and $\mathfrak{B}$. We can then prove

THEOREM 4. *The necessary and sufficient condition for the regular structure correspondence to establish a structure isomorphism*

$$(7)\qquad [A, B]/A \sim B/(A, B)$$

is that for every $\bar{A}$ and $\bar{B}$ defined by (5) *we have*

$$(8)\qquad \bar{A} = [A, (B, \bar{A})], \qquad \bar{B} = (B, [A, \bar{B}]).$$

Proof.[7] The relations (8) are obviously necessary and sufficient in order that the regular structure correspondence give a one-to-one correspondence between the subgroups of $\mathfrak{A}$ and $\mathfrak{B}$. Furthermore, when they are satisfied we find for $\bar{B}_1$ and $\bar{B}_2$ with the images $\bar{A}_1$ and $\bar{A}_2$

$$[\bar{B}_1, \bar{B}_2] \to [A, \bar{B}_1, \bar{B}_2] = [\bar{A}_1, \bar{A}_2],$$

$$(\bar{B}_1, \bar{B}_2) = (B, \bar{A}_1, \bar{A}_2) \to [A, (B, \bar{A}_1, \bar{A}_2)] = (\bar{A}_1, \bar{A}_2).$$

The conditions (8) are seen to hold if A has the following property: *For any C and D in $M = [A, B]$ we have*

$$(9)\qquad \begin{aligned} (C, [A, D]) &= [D, (A, C)], \qquad \text{when } C \geqq D, \\ (C, [A, D]) &= [A, (C, D)], \qquad \text{when } C \geqq A. \end{aligned}$$

In this case we shall say that A is *structure normal* in M.

Let us finally suppose that A and B are permutable so that the regular co-set correspondence gives a one-to-one correspondence between the co-sets of $\mathfrak{A}$ and $\mathfrak{B}$. We shall now make the stronger assumption that to every a in A and b in B there exists an exponent $n_{a,b}$ such that

$$(10)\qquad a \cdot b = b^{n_{a,b}} \cdot a',$$

where a' belongs to A. This condition is equivalent to saying that A shall be *permutable with all subgroups of B.*

We can then prove

THEOREM 5. *Let the group A be permutable with all subgroups of B. Then there exists a strong structure isomorphism*

$$(11)\qquad [A, B]/A \simeq B/(A, B),$$

and the regular co-set correspondence between the two quotients also gives the regular structure correspondence.

[7] This result has been derived for any structure in the paper by O. Ore, *On the theorem of Jordan-Hölder*, Transactions Amer. Math. Soc., vol. 41 (1937), pp. 266–275.

Proof. Any subgroup $\bar{B}$ is made up by co-sets $b_1D, b_2D, \cdots$, and to these correspond the co-sets $b_1A, b_2A, \cdots$. These co-sets must again form a group, since all elements in a product $b_1A \cdot b_2A$ are contained in co-sets of the form $b_1b_2^n \cdot A = b_3 \cdot A$. Hence we have $\bar{B} \to [A_1, \bar{B}]$, and it is seen immediately that the first relation (8) is satisfied. In the same manner any group $\bar{A}$ is made up by co-sets bA where the b form a group $\bar{B} = (\bar{A}, B)$ and also the second relation (8) is satisfied. Hence we have shown that the regular co-set correspondence gives a structure isomorphism identical with (5). The fact that (11) is a strong structure isomorphism follows from Theorem 2 or 3, since A is permutable with every $\bar{B}$.

It is also easily seen that in this case A is structure normal in $[A, B]$.

2. **Permutable groups.** The preceding results indicate the necessity of a study of the permutable groups.[8] We shall now give some of their simplest properties.

THEOREM 6. *If A and B are permutable groups, then*

$$(12)\qquad \begin{aligned} (C, [A, B]) &= [A, (C, B)], & C &\geqq A, \\ (D, [A, B]) &= [B, (A, D)], & D &\geqq B. \end{aligned}$$

Proof. To prove, for instance, the first of these relations, we observe that the right-hand side is always contained in the left-hand side for any groups. If now A and B are permutable, every element of $[A, B]$ has the form $a \cdot b$. If this element is to be at the same time an element c of C, we must have $c = a \cdot b$, where obviously b belongs to C, since $C \geqq A$.

Furthermore let us mention

LEMMA 1. *If A is permutable with B and C, then A is permutable with $[B, C]$.*

LEMMA 2. *A group is permutable with all its subgroups.*

LEMMA 3. *If $B \geqq B'$ and B' is permutable with A, then B' is permutable with (A, B).*

Proof. If d is an element of (A, B), then $b' \cdot d = c \cdot b_1'$, where c belongs to A and also to B.

LEMMA 4. *If $A \geqq A'$ and $B \geqq B'$ and A' and B' are permutable, (A', B) and (A, B') are permutable.*

The proof follows by two applications of the preceding lemma.

THEOREM 7. *Let $A \geqq A'$ and $B \geqq B'$. Furthermore let A' be permutable with (A, B) and (A, B'), while B' is permutable with (B, A) and (B, A'). Then there exists a one-to-one correspondence*

$$(13)\qquad \mathfrak{A} = [A', (A, B)]/[A', (A, B')] \rightleftarrows [B', (A, B)]/[B', (B, A')] = \mathfrak{B}$$

between the co-sets of these two quotients.

[8] Various properties of permutable groups have been given by E. Maillet, *Sur les groupes échangeables et les groupes décomposables*, Bulletin de la Société Math. de France, vol. 28 (1900), pp. 7–16.

Proof. We write

$$M = [A', (A, B')], \qquad N = (A, B).$$

From Lemmas 1 and 2 it follows that M and N are permutable, and from Theorem 6 we obtain

$$[M, N] = [A', (A, B')],$$

$$(M, N) = (A, B, [A', (A, B')]) = [(A, B'), (A', B)].$$

The application of Theorem 2 gives us the correspondence

$$\mathfrak{A} \rightleftarrows (A, B)/[(A, B'), (B, A')], \tag{14}$$

and the same correspondence is found for $\mathfrak{B}$.

Permutability of two groups is not a self-dual concept, since the structural relations (12) do not imply their duals. There exists, however, a theorem which may take the place of the dual of Theorem 7:

THEOREM 8. *Let $A \geqq A'$ and $B \geqq B'$ and let A and B both be permutable with $[A', B']$. Then there exists a correspondence*

$$\mathfrak{A}_1 = (A, [B, A'])/(A, [B', A']) \rightleftarrows (B, [A, B'])/(B, [A', B']) = \mathfrak{B}_1 . \tag{15}$$

Proof. In this case we write

$$M_1 = (A, [B, A']), \qquad N_1 = [A', B'],$$

and M_1 and N_1 are permutable according to Lemma 3. Theorem 6 gives

$$(M_1 , N_1) = (A, [B', A']),$$

$$[M_1 , N_1] = [A', B', (A, [B, A'])] = ([B, A'], [A, B']),$$

and from Theorem 2 one obtains

$$\mathfrak{A}_1 \rightleftarrows ([B, A'], [A, B'])/[A', B']. \tag{16}$$

THEOREM 9. *Let $A \geqq A'$ and $B \geqq B'$, where A' is permutable with B and B' and B' is permutable with A and A'. Then the correspondences* (13) *and* (15) *hold, and we have $\mathfrak{A} = \mathfrak{A}_1$ and $\mathfrak{B} = \mathfrak{B}_1$.*

Proof. It follows from Lemmas 1, 2 and 3 that in this case the conditions for Theorems 7 and 8 are satisfied. The equality of the quotients is obtained from Theorem 6. From the correspondences (14) and (16) we deduce the following self-dual relation:

THEOREM 10. *Let $A \geqq A'$ and $B \geqq B'$, where A is permutable with B and B', while B' is permutable with A and A'. Then*

$$(A, B)/(A, B, [A', B']) \simeq [A', B', (A, B)]/[A', B']. \tag{17}$$

One finds, namely, from Theorem 6,

$$[(A, B'), (B, A')] = (A, B, [A', B']),$$

$$([A, B'], [B, A']) = [A', B', (A, B)].$$

There are various other interesting problems connected with permutable groups. One of the main problems is the determination of all permutable groups. It may be formulated as follows. *Let A and B be given groups. Find all groups $M = [\bar{A}, \bar{B}]$, where $\bar{A}$ and $\bar{B}$ are permutable groups isomorphic to A and B respectively.*

Another problem is the relation of the permutable groups to their structural properties. This problem may be formulated as a converse of Theorem 6: *when do the relations* (12) *imply that A and B are permutable groups?*

We shall next determine when the correspondences we have derived may be replaced by strong structure isomorphisms. According to Theorem 5 we shall then have to consider groups which are permutable with all subgroups of other groups. For such groups we have

LEMMA 5. *If B and C are permutable with all subgroups of A, then $[B, C]$ has the same property.*

LEMMA 6. *If $A \geqq A'$, then A is permutable with all subgroups of A'.*

LEMMA 7. *If $A \geqq A'$ and B is permutable with all subgroups of A', then (A, B) has the same property.*

Proof. Let a' be an element in A' and d an element in (A, B). Then $d \cdot a' = a''' \cdot b$, where b belongs both to A and B.

From these remarks we obtain

THEOREM 11. *Let $A \geqq A'$ and $B \geqq B'$, where A' and B' are permutable with all subgroups of (A, B). Then*

$$[A', (A, B)]/[A', (A, B')] \simeq [B', (A, B)]/[B', (B, A')].$$

Proof. In this case the conditions of Theorem 7 are satisfied. Furthermore, M is permutable with all subgroups of N according to Lemmas 5 and 7. Hence we conclude from Theorem 5 the strong structure isomorphism

$$\mathfrak{A} \simeq (A, B)/[(A, B'), (B, A')],$$

and similarly for $\mathfrak{B}$. Corresponding to Theorem 8 we have

THEOREM 12. *Let $A \geqq A'$ and $B \geqq B'$, where $[A', B']$ is permutable with every subgroup of A and B. Then*

$$(A, [B, A'])/(A, [B', A']) \simeq (B, [A, B'])/(B, [A', B']).$$

Corresponding to Theorem 10 we have

THEOREM 13. *Let $A \geqq A'$ and $B \geqq B'$, where A' and B' are permutable with every subgroup of A and B. Then*

$$(A, B)/(A, B, [A', B']) \simeq [A', B', (A, B)]/[A', B'].$$

Let us observe finally that if A and B are groups such that A is permutable with every subgroup of B and B permutable with every subgroup of A, the rela-

tions (12) obviously must hold. In this case it is seen, however, that also the dual relations

$$(18)\qquad \begin{aligned} [C, (A, B)] &= (A, [B, C]), & C &\leqq A, \\ [D, (A, B)] &= (B, [A, D]), & D &\leqq B \end{aligned}$$

are fulfilled. Again the interesting problem arises whether the existence of the relations (12) and (18) is sufficient to conclude that A is permutable with every subgroup of B and B permutes with every subgroup of A.

3. **Quasi-normal subgroups.** We shall now introduce a new concept.

A subgroup A of G is said to be quasi-normal when it is permutable with every subgroup of G.

This condition may also be stated as follows. For each g in G and a in A there exists an exponent $n_{a,g}$ such that

$$(19)\qquad a \cdot g = g^{n_{a,g}} \cdot a'.$$

The quasi-normal subgroups obviously generalize the ordinary normal subgroups. In that case we have $n_{a,g} = 1$ for all a and g. The condition (19) may also be stated in the more symmetric form

$$a \cdot g^{n} = g^{n'} \cdot a'.$$

The quasi-normal subgroups have several properties in common with ordinary normal subgroups. Let us mention first that from Theorem 6 follows

THEOREM 14. *If A is a quasi-normal subgroup and B and C arbitrary subgroups, then*

$$\begin{aligned} (B, [A, C]) &= [C, (A, B)] & &\textit{when } B \geqq C, \\ (B, [A, C]) &= [A, (B, C)] & &\textit{when } B \geqq A. \end{aligned}$$

It is obvious that when A and B are quasi-normal in G, $[A, B]$ has the same property.

THEOREM 15. *If A is quasi-normal in G and B is any subgroup of G, then (A, B) is quasi-normal in B.*

Proof. Let d be any element of (A, B) and b any element of B. Then $d \cdot b = b^{n} \cdot a$, where a must belong both to A and B. From Theorem 5 follows immediately

THEOREM 16. *When A is quasi-normal in $[A, B]$, then (A, B) is quasi-normal in B, and we have the strong structure isomorphism*

$$(20)\qquad [A, B]/A \simeq B/(A, B).$$

To prove the next theorem we shall need

LEMMA 8. *Let A be quasi-normal in G. When g is any element in G and a_0 some given element in A, we can always write*

$$g^{n} = (ga_0)^{m} \cdot a.$$

Proof. We write $g = ga_0 \cdot a_0^{-1}$ and apply (19) to the various factors in the product g^n.

THEOREM 17. *In the structure isomorphism* (20) *any quasi-normal group* $\bar{A}$ *in* $[A, B]$ *containing* A *corresponds to a quasi-normal group* $\bar{B}$ *in* B *containing* (A, B), *and conversely.*

Proof. The first part of the theorem follows from Theorem 15, since $\bar{B} = (\bar{A}, B)$. To prove the converse we consider the relation

$$\bar{b} \cdot b = b^n \cdot \bar{b}_1, \tag{21}$$

which holds for all elements in $\bar{B}$ and B. The elements of $[A, B]$ have the form $g = b \cdot a$, while the elements of $\bar{A}$ are of the form $a \cdot \bar{b}$. By means of Lemma 8 we obtain from (21) $\bar{b} \cdot g = g^n \cdot \bar{a}_1$, and when this relation is multiplied on the left by an element in A we obtain $\bar{a} \cdot g = g^m \cdot \bar{a}_2$ for arbitrary $\bar{a}$ in $\bar{A}$.

Theorem 17 gives us the ordinary lemma that if A and B are maximal quasi-normal groups in G, then (A, B) is maximal quasi-normal in A and B.

THEOREM 18. *When* A' *is quasi-normal in* A *and* B' *quasi-normal in* B, *then* $[A', (A, B')]$ *is quasi-normal in* $[A', (A, B)]$ *and similarly* $[B', (A', B)]$ *is quasi-normal in* $[B', (A, B)]$, *and there exists the strong structure isomorphism*

$$[A', (A, B)]/[A', (A, B')] \simeq [B', (A, B)]/[B', (B, A')]. \tag{22}$$

Proof. The isomorphism (22) is a consequence of Theorem 11. Let $l = d \cdot a'$ be an element of $[(A, B), A']$ and $l' = d' \cdot a'$ be an element in $[(A, B'), A']$. Since (A, B') is quasi-normal in (A, B), we find

$$l' \cdot l = d' \cdot a_1' \cdot d \cdot a' = d' \cdot d^n \cdot a_2' = d^m \cdot l_1'.$$

According to Lemma 8 we can write $d^m = l^k \cdot a_3'$, so that we finally obtain $l' \cdot l = l^k \cdot l_2'$.

Theorem 18 represents the complete analogue of the lemma of Zassenhaus for ordinary normal subgroups.

Let A' *be normal in* A *and* B' *normal in* B. *Then* $[A', (A, B')]$ *is normal in* $[A', (A, B)]$ *and* $[B', (A', B)]$ *is normal in* $[B', (A, B)]$, *and there exists the ordinary isomorphism between the quotient groups*

$$[A', (A, B)]/[A', (A, B')] \cong [B', (A, B)]/[B', (A', B)].$$

4. **Quasi-normal subgroups of the symmetric and alternating groups.** In connection with the quasi-normal groups one may construct groups which in many ways are analogous to those ordinarily defined in connection with normal subgroups. Among these we shall mention the *quasi-center*. Let c be an element of G with the property that the cyclic group it generates is permutable with all subgroups of G. This is equivalent to saying that for any element g of G and any pair of exponents n and m we have $c^n \cdot g^m = g^{m'} \cdot c^{n'}$. The group generated by these elements c we shall call the *quasi-center* of G. The quasi-

center contains the ordinary *center* and also the *nucleus* introduced by Baer.[9] From the definition of the quasi-center follows

THEOREM 19. *The quasi-center is a characteristic subgroup.*

A simple consideration shows that the symmetric group Σ_n has a quasi-center C_n equal to the unit element except in the case $n = 2$, where $\Sigma_2 = C_2$, and $n = 3$, where $C_3 = A_3$ is the alternating group. The quasi-center of the alternating group is the unit element except for $n = 3$.

It may be of some interest to determine the quasi-normal subgroups of the symmetric group Σ_n and the alternating group A_n. The case of the symmetric group may be solved directly by the following

LEMMA 9. *Let H be quasi-normal in G. If h is an element of H and t an element of order 2 in G, then H contains the transform tht^{-1} and the commutator $(t, h) = tht^{-1}h^{-1}$.*

Proof. Since we can suppose that t does not belong to H, we have $th = h'\cdot t$. Since every substitution is the product of transpositions this implies

THEOREM 20. *Any quasi-normal subgroup of the symmetric group is normal.*

The proof of the theorem that the alternating group contains only the trivial quasi-normal groups A_n and E is necessarily more complicated, since it implies the fact A_n is simple.

We first show

LEMMA 10. *If a quasi-normal subgroup H of A_n contains a cycle with three elements, then $H = A_n$.*

Proof. We may suppose $n \geqq 4$. If H contains the cycle $h = (1, 2, 3)$, according to Lemma 9 it contains

$$aha^{-1} = (\alpha, 3, 2), \qquad a = (1, \alpha)(2, 3), \qquad \alpha > 3,$$

and hence H contains all cycles of order 3 with two fixed elements. We shall also need

LEMMA 11. *Let $n > 4$ and let H be a quasi-normal subgroup of A_n. If H contains a substitution which is the product of two transpositions without common elements, then $H = A_n$.*

Proof. Let $h = (1, 2)(3, 4)$ belong to H. The commutator

$$(a, h) = (3, 5, 4), \qquad a = (1, 2)(3, 5)$$

then also belongs to H according to Lemma 9.

We shall now proceed to the actual determination of the quasi-normal subgroups of A_n. We suppose $n > 4$ and write the substitutions of H as the product of cycles without common letters. We consider the cases

I. H contains a substitution $h = (1, 2, \cdots, \alpha)$ with at least one cycle with $\alpha > 4$ letters. We put $a = (1, 3)(2, 4)$ and find the commutator $(a, h) = (1, 3, 5)$.

[9] R. Baer, *Der Kern, eine charakteristische Untergruppe*, Compositio Mathematica, vol. 1 (1934), pp. 254–283.

II. H contains a substitution $h = (1, 2, 3, 4) \cdots$ with at least one cycle of order 4. Then we put $a = (1, 4)(2, 3)$ and find $(a, h) = (1, 3)(2, 4)$.

III. H contains substitutions h with a cycle of order 3. If h contains only one such cycle while the other cycles are transpositions, then h^2 is a cycle of order 3. Hence we may suppose $h = (1, 2, 3)(4, 5, 6) \cdots$ and for $a = (1, 4)(2, 5)$, we find $(a, h) = (1, 4)(3, 6)$.

IV. All substitutions in H are the product of transpositions. According to Lemma 11 we can suppose that h contains at least three transpositions $h = (1, 2)(3, 4)(5, 6) \cdots$ and for $a = (1, 4)(2, 5)$ we find $(a, h) = (1, 6, 3)(2, 4, 5)$ against assumption.

THEOREM 21. *The alternating group contains no quasi-normal subgroup for* $n \neq 4$.

For $n = 4$ we have the well-known exception.

Chapter 3. Extensions of the theorem of Jordan-Hölder

1. **Refinements of chains.** One of the main applications of the preceding theory is to obtain extensions of the theorem of Jordan-Hölder and its generalization by Schreier-Zassenhaus.[10]

In the following we shall consider two fixed groups $A > B$ in the given group G. Let

$$(1) \qquad A = B_0 > B_1 > \cdots > B_r = B,$$

$$(2) \qquad A = C_0 > C_1 > \cdots > C_s = B$$

be two chains of arbitrary subgroups between A and B. We shall denote these chains by $\{B_i\}$ and $\{C_j\}$ respectively. The existence of the chains (1) and (2) may also be considered as a factorization of the corresponding quotient $\mathfrak{A} = A/B$ in the sense defined in Chapter 1,

$$(3) \qquad \mathfrak{A} = \mathfrak{B}_1 \times \mathfrak{B}_2 \times \cdots \times \mathfrak{B}_r = \mathfrak{C}_1 \times \mathfrak{C}_2 \times \cdots \times \mathfrak{C}_s,$$

where

$$(4) \qquad \mathfrak{B}_i = B_{i-1}/B_i, \qquad \mathfrak{C}_j = C_{j-1}/C_j.$$

Any new chains obtained from (1) and (2) by intercalating new terms will be called a *refinement* of the given chains. The two refinements consisting both of $r \cdot s$ subgroups

$$(5) \qquad B_{i,j} = [B_i, (C_j, B_{i-1})], \qquad C_{k,l} = [C_k, (B_l, C_{k-1})]$$

we shall call the (*lower*) *refinement of* $\{B_i\}$ *with respect to* $\{C_j\}$ and of $\{C_j\}$ with respect to $\{B_i\}$, respectively. It may also be considered as a further factorization of (3):

[10] O. Schreier, *Über den Jordan-Hölderschen Satz*, Abh. Math. Sem. Hamburg, vol. 6 (1928), pp. 300–302; H. Zassenhaus, *Zum Satz von Jordan-Hölder-Schreier*, Abh. Math. Sem. Hamburg, vol. 10 (1934), pp. 106–108.

$$\mathfrak{B}_i = \mathfrak{B}_{i,1} \times \cdots \times \mathfrak{B}_{i,s}, \qquad \mathfrak{C}_j = \mathfrak{C}_{j,1} \times \cdots \times \mathfrak{C}_{j,r}, \tag{6}$$

where

$$\begin{aligned} \mathfrak{B}_{i,j} &= [B_i, (C_{j-1}, B_{i-1})]/[B_i, (C_j, B_{i-1})], \\ \mathfrak{C}_{k,l} &= [C_k, (C_{k-1}, B_{l-1})]/[C_k, (C_{k-1}, B_l)]. \end{aligned} \tag{7}$$

There also exists a dual refinement to (5):

$$B'_{i,j} = (B_{i-1}, [B_i, C_j]), \qquad C'_{k,l} = (C_k, [B_l, C_{k-1}]), \tag{8}$$

which we shall call the upper refinement of $\{B_i\}$ with respect to $\{C_j\}$ and of $\{C_j\}$ with respect to $\{B_i\}$.

Obviously we always have

$$B'_{i,j} \geqq B_{i,j}, \qquad C'_{k,l} \geqq C_{k,l}.$$

Corresponding to (6) we have the factorizations

$$\mathfrak{B}_i = \mathfrak{B}'_{i,1} \times \cdots \times \mathfrak{B}'_{i,s}, \qquad \mathfrak{C}_j = \mathfrak{C}'_{j,1} \times \cdots \times \mathfrak{C}'_{j,r}, \tag{9}$$

where

$$\begin{aligned} \mathfrak{B}'_{i,j} &= (B_{i-1}, [B_i, C_{j-1}])/(B_{i-1}, [B_i, C_j]), \\ \mathfrak{C}'_{k,l} &= (C_{k-1}, [C_k, B_{l-1}])/(C_{k-1}, [C_k, B_l]). \end{aligned} \tag{10}$$

2. Cross-cut permutable chains. We shall now say that two chains (1) and (2) are *cross-cut permutable* if they have the following property:

For each i the group B_i shall be permutable with all groups

$$(B_{i-1}, C_j) \qquad (j = 0, 1, \cdots, s), \tag{11}$$

and for each j the group C_j shall be permutable with

$$(C_{j-1}, B_i) \qquad (i = 0, 1, \cdots, r). \tag{12}$$

Let us observe that cross-cut permutability refers only to certain properties of a group B_i with respect to certain subgroups of the preceding B_{i-1}, and similarly for C_j with respect to C_{j-1}.

The main theorem on cross-cut permutable chains will now be proved.

THEOREM 1. *Let $\{B_i\}$ and $\{C_j\}$ be two arbitrary cross-cut permutable chains connecting two groups A and B. These chains can then be refined into new cross-cut permutable chains with quotients corresponding in such a manner that for corresponding quotients there is a one-to-one correspondence of their co-sets.*

Proof. We refine the two chains by the lower refinements (5). Both the refined chains contain $r \cdot s$ terms, and it follows from Theorem 7, Chapter 2, that $\mathfrak{B}_{i,j} \rightleftarrows \mathfrak{C}_{j,i}$.

The proof of the fact that the new chains are cross-cut permutable is more complicated. It may be done directly by calculating the groups corresponding to (11) for the new chains. We shall, however, apply another method which yields more information about cross-cut permutable chains. We first deduce

LEMMA 1. *The chains $\{B_{i,j}\}$ and $\{C_k\}$ are cross-cut permutable.*

Proof. We shall have to show that $B_{i,j}$ is permutable with every $(B_{i,j-1}, C_k)$. Since B_{i-1} contains $B_{i,j-1}$ we can write

$$(13) \qquad (B_{i,j-1}, C_k) = D_{i,j,k} = (C_k, B_{i-1}, [B_i, (C_{j-1}, B_{i-1})]).$$

If now $k \geqq j$, then (C_k, B_{i-1}) is contained in (C_{j-1}, B_{i-1}), and we find

$$(14) \qquad D_{i,j,k} = (C_k, B_{i-1}) \qquad (k \geqq j).$$

When $k \leqq j - 1$, we apply Theorem 6, Chapter 2, to (13) and obtain

$$(15) \qquad D_{i,j,k} = [(C_{j-1}, B_{i-1}), (B_i, C_k)] \qquad (k \leqq j - 1).$$

To show that $B_{i,j}$ is permutable with $D_{i,j,k}$, it is sufficient to show according to Lemma 1, Chapter 2, that B_i and $B_{i-1,j}$ separately are permutable with $D_{i,j,k}$. It is obvious from the given conditions that B_i is permutable both with (14) and (15). Furthermore, (B_{i-1}, C_j) is permutable with (14) according to Lemma 2 and with (15) according to Lemmas 2 and 3, Chapter 2.

Finally one proves in a similar manner that C_k is permutable with the cross-cuts $D_{i,j,k-1} = (B_{i,j}, C_{k-1})$.

LEMMA 2. *The chain $\{B_{i,j}\}$ is unchanged when refined with respect to $\{C_j\}$.*

Proof. The terms of the second refinement are $B_{i,j,k} = [B_{i,j}, (C_k, B_{i,j-1})]$. When the expressions (14) and (15) are substituted here, one obtains respectively

$$B_{i,j,k} = [B_i, (B_{i-1}, C_j), (B_{i-1}, C_k)] = [B_i, (B_{i-1}, C_j)] = B_{i,j},$$

$$B_{i,j,k} = [B_i, (B_{i-1}, C_j), (C_{j-1}, B_{i-1}), (B_i, C_k)] = B_{i,j-1}.$$

LEMMA 3. *The refinement of $\{C_j\}$ with respect to $\{B_i\}$ is equal to its refinement with respect to $\{B_{i,j}\}$.*

Proof. We find $C_{i,j,k} = [C_k, (C_{k-1}, B_{i,j})]$, and again the substitution of (14) and (15) gives

$$C_{i,j,k} = C_{k,i-1} \qquad (k \geqq j)$$

$$C_{i,j,k} = C_{k,i} \qquad (k \leqq j - 1).$$

From this lemma, together with Lemma 1, it follows that the chains $\{B_{i,j}\}$ and $\{C_{k,l}\}$ are cross-cut permutable and the proof of Theorem 1 is completed.

The preceding results also show that repeated refinements of cross-cut permutable chains give no new chains.

Theorem 1 is obviously a generalization of the theorem of Schreier-Zassenhaus. Through specialization one can also obtain extensions of the theorem of Jordan-Hölder. We shall say that $\{B_i\}$ and $\{C_j\}$ are *maximal cross-cut permutable chains* when there exists no group between B_{i-1} and B_i which is permutable with all cross-cuts (11) and similarly for C_{j-1} and C_j. We then have

THEOREM 2. *If there exist two maximal cross-cut permutable chains between the two groups A and B, both chains have the same number of terms and the quotients*

of the chains correspond in such a manner that for corresponding quotients there is a one-to-one correspondence between their co-sets.

When the index of B in A is finite, Theorem 2 shows that the indices of the two maximal cross-cut permutable chains are the same in some order.

There exists for this theory a dual theory which one obtains by considering the upper refinements (8). We shall say that two chains $\{B_i\}$ and $\{C_j\}$ are *union permutable* when they have the following property:

For all i the group B_{i-1} is permutable with

$$[B_i, C_j] \qquad (j = 0, 1, \cdots, s),$$

and for all j C_{j-1} is permutable with

$$[C_j, B_i] \qquad (i = 0, 1, \cdots, r).$$

For such union permutable chains all the theorems derived for cross-cut permutable chains will hold when the lower refinements $\{B_{i,j}\}$ and $\{C_{k,l}\}$ are replaced by the upper refinements $\{B'_{i,j}\}$ and $\{C'_{k,l}\}$. In Theorem 2 we must replace maximal chains by minimal chains.

3. **Permutable chains.** We have already observed that Theorems 1 and 2 correspond to the theorems of Schreier-Zassenhaus and Jordan-Hölder. More specifically they correspond to the case of composition series where one considers chains in which every term is a normal subgroup of the preceding, because cross-cut permutability refers only to properties of a group B_i with respect to certain subgroups of the preceding B_{i-1}. Let us now show how one can obtain analogues of the theorems on principal series where one deals with normal subgroups of the full group A.

We shall say that the two chains $\{B_i\}$ and $\{C_j\}$ are *permutable* if every B_i is permutable with every C_j. We can then prove

THEOREM 3. *Let $\{B_i\}$ and $\{C_j\}$ be two permutable chains connecting A and B. These chains may be refined into new permutable chains with their quotients corresponding in such a manner that for corresponding quotients there is a one-to-one correspondence between their co-sets.*

Proof. It follows from Theorem 9, Chapter 2, that in this case the upper and the lower refinements (5) and (8) are identical. Furthermore, we have $\mathfrak{B}_{i,j} \rightleftarrows \mathfrak{C}_{j,i}$. The fact that the refined chains are again permutable may be derived from Lemmas 1 to 4 in Chapter 2.

Through specialization of Theorem 3 we can again obtain an analogue to the Jordan-Hölder theorem. We shall say that two permutable chains are *maximal* when no further term may be intercalated such that the resulting chains are still permutable.

THEOREM 4. *Any two maximal permutable chains between two groups $A > B$ will have the same number of terms and their quotients correspond in such a manner that for corresponding quotients there exists a one-to-one correspondence between their co-sets.*

For finite groups this theorem, in the slightly weaker form when the indices of the two chains are the same in some order, is due to Maillet.[11]

4. **Quasi-normal chains.** The next natural step is to seek structure isomorphism for the corresponding quotients in the two chains. We shall say that a chain is *quasi-normal* when each term is quasi-normal in the preceding. The main theorem is then

THEOREM 5. *Let* $\{B_i\}$ *and* $\{C_j\}$ *be two quasi-normal chains connecting* A *and* B. *These chains may be refined into new quasi-normal chains in such a way that there exists a correspondence between their quotients, with corresponding quotients having strongly structure isomorphic quotient systems.*

Proof. We apply the lower refinements (5) and all statements of Theorem 5 are consequences of Theorem 18, Chapter 2.

Theorem 5 is a complete analogue of the theorem of Schreier-Zassenhaus. The analogue of the theorem of Jordan-Hölder may be expressed as follows:

THEOREM 6. *If there exist two maximal quasi-normal chains between* A *and* B, *both contain the same number of terms and the quotient systems are strongly structure isomorphic in some order.*

A maximal quasi-normal chain is of course a quasi-normal chain in which no further quasi-normal terms may be intercalated.

One might finally want to obtain an analogue to the theorem on principal chains by considering chains in which every term is quasi-normal in the full group A. This is, however, the only point at which our theory differs from the ordinary normal theory. In this case we cannot prove that the refined chains are again formed by quasi-normal groups in A. This is due to the fact that the quasi-normal groups in A do not form a structure, since the cross-cut of two quasi-normal groups need not be quasi-normal in A.

Chapter 4. Normal subgroups

1. **Similarity.** Let us now turn to the properties of normal subgroups. The normal subgroups of a group G obviously form a structure. Since any two normal subgroups are permutable we also have:

Any three normal subgroups $A \geqq B$ *and* C *satisfy the Dedekind axiom*

$$(A, [B, C]) = [B, (A, C)]. \tag{1}$$

A structure satisfying the Dedekind axiom I have called a *Dedekind structure.*[12] The general theory of Dedekind structures may now be applied to the normal subgroups. Since the Dedekind axiom (1) is found to be a self-dual condition,

[11] E. Maillet, *Sur de nouvelles analogies entre la théorie des groupes de substitutions et celles des groupes finis continus de transformations de Lie, première note: Sur des suites remarquables de sous-groupes d'un groupe de substitutions*, Journal de Math., (5), vol. 7 (1901). pp. 13–82.

[12] *Foundations* I, Chapter 1.

our principle of duality holds to its full extent also for normal subgroups. One may generalize the theory of normal subgroups in the ordinary way by considering groups with operators. An operator T in G has the property of making correspond to each element a an element a^T such that

$$(a \cdot b)^T = a^T \cdot b^T.$$

A subgroup A is invariant with respect to T if

$$A^T \leqq A.$$

All groups which are invariant under a certain system of operators

$$\mathfrak{T} = T_1, T_2, \cdots$$

also form a structure, and this structure is a Dedekind structure in case the operators include the set of all inner automorphisms of G.

In the structure of normal subgroups we introduce quotients and these are as usual associated with a quotient group. Such quotients can be transformed according to the rules given in Chapter 1. Let

$$\mathfrak{A} = A/B, \qquad \mathfrak{C} = C/B \tag{2}$$

be two quotients with the same denominator B, whereby we assume as always in the following that the subgroups are considered as normal in G. We suppose that $\mathfrak{A}$ and $\mathfrak{C}$ are relatively prime, i.e., $B = (A, C)$. According to the ordinary law of isomorphism the right-hand similarity transformation

$$\mathfrak{C}\mathfrak{A}\mathfrak{C}^{-1} = [A, C]/C \tag{3}$$

then gives a new quotient group which is isomorphic to $\mathfrak{A}$ in the usual sense. In the more general case where $\mathfrak{A}$ and $\mathfrak{C}$ have a greatest common factor $\mathfrak{D}_1$,

$$\mathfrak{A} = \mathfrak{A}_1 \times \mathfrak{D}_1, \qquad \mathfrak{C} = \mathfrak{C}_1 \times \mathfrak{D}_1,$$

we have according to Lemma 4, Chapter 1,

$$\mathfrak{C}\mathfrak{A}\mathfrak{C}^{-1} = \mathfrak{C}_1\mathfrak{A}_1\mathfrak{C}_1^{-1},$$

so that the transform is isomorphic to the left-hand factor $\mathfrak{A}_1$ of $\mathfrak{A}$.

In the same manner, if

$$\mathfrak{A} = A/B, \qquad \mathfrak{D} = A/D, \qquad A = [B, D] \tag{4}$$

are two left-hand relatively prime quotients with the same numerator, the left-hand similarity transformation

$$\mathfrak{D}^{-1}\mathfrak{A}\mathfrak{D} = D/(B, D) \tag{5}$$

gives a quotient group isomorphic with $\mathfrak{A}$. Again if $\mathfrak{A}$ and $\mathfrak{D}$ in (4) have a common l.h. factor, the transform (5) is isomorphic to a r.h. factor of $\mathfrak{A}$.

Two quotients $\mathfrak{A}$ and $\mathfrak{B}$ shall be said to be *similar* if one can be obtained from the other through a series of r.h. and l.h. similarity transformations. Similar

quotients are isomorphic. Similarity is a special form of isomorphism and it follows from the general theory of Dedekind structures that it is the form of isomorphism which occurs in the formulation of the decomposition theorems for algebraic systems.

We shall say that when $\mathfrak{A}$ and $\mathfrak{C}$ are r.h. relatively prime quotients with the same denominator defined by (2), then

$$\mathfrak{M} = [\mathfrak{A}, \mathfrak{C}] \tag{6}$$

is the *direct union* of $\mathfrak{A}$ and $\mathfrak{C}$. This corresponds of course to the ordinary direct product. In the same manner one defines the left-hand direct union. It follows from Theorem 6, Chapter 1, that to any r.h. direct union (6) corresponds a l.h. direct union $\mathfrak{M} = [\mathfrak{A}', \mathfrak{C}']_l$, where $\mathfrak{A}'$ and $\mathfrak{C}'$ are similar to $\mathfrak{A}$ and $\mathfrak{C}$. This result may be extended to an arbitrary number of quotients.

2. **Direct similarity.** We shall now study a special type of similarity. Let $\mathfrak{A}$ and $\mathfrak{B}$ be quotients with the same denominator. If there exists a third quotient $\mathfrak{C}$ with the same denominator and relatively prime to both $\mathfrak{A}$ and $\mathfrak{B}$ such that

$$\mathfrak{H} = [\mathfrak{A}, \mathfrak{C}] = [\mathfrak{B}, \mathfrak{C}], \qquad (\mathfrak{A}, \mathfrak{C}) = (\mathfrak{B}, \mathfrak{C}) = \mathfrak{C}_0, \tag{7}$$

we shall say that $\mathfrak{A}$ and $\mathfrak{B}$ are *directly similar in* $\mathfrak{H}$. The relation (7) may also be written

$$\mathfrak{C}\mathfrak{A}\mathfrak{C}^{-1} = \mathfrak{C}\mathfrak{B}\mathfrak{C}^{-1}.$$

When $\mathfrak{A}$ and $\mathfrak{B}$ are directly similar in some $\mathfrak{H}$, they are also directly similar in $[\mathfrak{A}, \mathfrak{B}]$. One obtains therefore by taking the cross-cut of $\mathfrak{H}$ with $[\mathfrak{A}, \mathfrak{B}]$

$$[\mathfrak{A}, \mathfrak{C}_1] = [\mathfrak{B}, \mathfrak{C}_1] = [\mathfrak{A}, \mathfrak{B}], \qquad (\mathfrak{A}, \mathfrak{C}_1) = (\mathfrak{B}, \mathfrak{C}_1) = \mathfrak{C}_0, \tag{8}$$

where, according to the Dedekind axiom,

$$\mathfrak{C}_1 = (\mathfrak{C}, [\mathfrak{A}, \mathfrak{B}]).$$

The isomorphism between $\mathfrak{A}$ and $\mathfrak{B}$ defined by (8) is such that to any a in $\mathfrak{A}$ there corresponds a unique b in $\mathfrak{B}$ with the property that

$$a \cdot c_1 = b \cdot c_2, \tag{9}$$

where c_1 and c_2 belong to $\mathfrak{C}_1$. The direct decompositions (8) show, however, that the elements of $\mathfrak{C}$ are permutable with the elements of both $\mathfrak{A}$ and $\mathfrak{B}$; hence $\mathfrak{C}$ belongs to the center of $[\mathfrak{A}, \mathfrak{B}]$. The correspondence (9) may then be written $a = b \cdot c$ where c belongs to the center, i.e., we have a *central isomorphism* between $\mathfrak{A}$ and $\mathfrak{B}$ in the ordinary sense in which this term is used in group theory.

THEOREM 1. *If two quotients $\mathfrak{A}$ and $\mathfrak{B}$ are directly similar, they are centrally isomorphic in the union* $[\mathfrak{A}, \mathfrak{B}]$.

The following theorem is also a simple consequence of the preceding remarks:

Theorem 2. *Let $\mathfrak{A}$, $\mathfrak{B}$ and $\mathfrak{C}$ be quotient groups with the same denominator. The relations*

$$(10) \qquad \mathfrak{M} = [\mathfrak{A}, \mathfrak{B}] = [\mathfrak{A}, \mathfrak{C}] = [\mathfrak{B}, \mathfrak{C}], \qquad (\mathfrak{A}, \mathfrak{C}) = (\mathfrak{B}, \mathfrak{C}) = \mathfrak{E}_0$$

imply that $\mathfrak{C}$ is an Abelian group belonging to the center of $\mathfrak{M}$. The relations

$$(11) \qquad \mathfrak{M} = [\mathfrak{A}, \mathfrak{C}] = [\mathfrak{B}, \mathfrak{C}], \qquad (\mathfrak{A}, \mathfrak{B}) = (\mathfrak{B}, \mathfrak{C}) = (\mathfrak{A}, \mathfrak{C}) = \mathfrak{E}_0$$

imply that $\mathfrak{A}$ and $\mathfrak{B}$ are Abelian groups. Finally, the self-dual set of relations

$$(12) \quad \mathfrak{M} = [\mathfrak{A}, \mathfrak{B}] = [\mathfrak{A}, \mathfrak{C}] = [\mathfrak{B}, \mathfrak{C}], \qquad (\mathfrak{A}, \mathfrak{B}) = (\mathfrak{A}, \mathfrak{C}) = (\mathfrak{B}, \mathfrak{C}) = \mathfrak{E}_0$$

implies that $\mathfrak{M}$ is an Abelian group.

Let us also observe that any relation (7) implies that $\mathfrak{A}$ and $\mathfrak{B}$ have Abelian left-hand factors which are directly similar. This follows from (7) by division with $(\mathfrak{A}, \mathfrak{B})$ and application of (12).

The dual of Theorem 1 is of considerable interest. Let $\mathfrak{A}_1$, $\mathfrak{B}_1$ and $\mathfrak{C}_1$ be quotients with the same numerator such that

$$(13) \qquad \mathfrak{M} = [\mathfrak{A}_1, \mathfrak{C}_1]_l = [\mathfrak{B}_1, \mathfrak{C}_1]_l, \qquad (\mathfrak{A}_1, \mathfrak{C}_1)_l = (\mathfrak{B}_1, \mathfrak{C}_1)_l = \mathfrak{E}_0 .$$

We shall then say that $\mathfrak{A}_1$ and $\mathfrak{B}_1$ are *left-hand directly similar*. Again it is no limitation to assume that in addition to (13) also the relation

$$(14) \qquad \mathfrak{M} = [\mathfrak{A}_1, \mathfrak{B}_1]_l$$

holds. In order to analyze the content of these relations, we write

$$\mathfrak{M} = M/D, \qquad \mathfrak{A}_1 = M/A, \qquad \mathfrak{B}_1 = M/B, \qquad \mathfrak{C}_1 = M/C,$$

and find that the conditions (13) and (14) are equivalent to

$$M = [A, C] = [B, C], \qquad (A, B) = (A, C) = (B, C) = D.$$

These relations can, however, be written as

$$(15) \qquad [\mathfrak{A}, \mathfrak{C}] = [\mathfrak{B}, \mathfrak{C}], \qquad (\mathfrak{A}, \mathfrak{B}) = (\mathfrak{A}, \mathfrak{C}) = (\mathfrak{B}, \mathfrak{C}) = \mathfrak{E}_0 ,$$

where we have introduced the quotients

$$\mathfrak{A} = A/D, \qquad \mathfrak{B} = B/D, \qquad \mathfrak{C} = C/D.$$

According to Theorem 2 the relations (15) imply that $\mathfrak{A}$ and $\mathfrak{B}$ are Abelian groups. But since

$$\mathfrak{C}\mathfrak{A}\mathfrak{C}^{-1} = \mathfrak{C}\mathfrak{B}\mathfrak{C}^{-1} = \mathfrak{C}_1 ,$$

we also obtain that $\mathfrak{C}_1$ is Abelian. This in turn means that since

$$\mathfrak{M} = \mathfrak{C}_1 \times \mathfrak{C},$$

the group $\mathfrak{C}$ contains the *commutator group* of $\mathfrak{M}$. To obtain a completely dualistic formulation of these results let us define the *anti-center* of a group as the quotient group with respect to the commutator group. We then have

THEOREM 3. *Let $\mathfrak{A}_1$ and $\mathfrak{B}_1$ be left-hand directly similar quotient groups. Then we have the relations*

$$\mathfrak{M} = [\mathfrak{A}_1, \mathfrak{B}_1]_l = [\mathfrak{A}_1, \mathfrak{C}_1]_l, \qquad (\mathfrak{A}_1, \mathfrak{C}_1)_l = (\mathfrak{B}_1, \mathfrak{C}_1)_l = \mathfrak{C}_0,$$

where $\mathfrak{C}_1$ is an Abelian group which is a left-hand factor of the anti-center of $\mathfrak{M}$.

This theorem is the dual of Theorem 1. It shows the duality between the center and the anti-center of a group. The existence of such a duality was already observed by Speiser[13] in connection with the direct product decompositions of a group. Here it is explained as an instance of our general principle of duality.

3. **Properties of three normal subgroups.** Since the Dedekind axiom is satisfied for normal subgroups, any three normal subgroups will generate by cross-cuts and unions a special Dedekind structure containing in general 28 normal subgroups. The discussion of this structure leads us to a set of relations between three normal subgroups which we have analyzed in *Foundations* I, and which we shall not repeat here. We shall only make one application to obtain a particular theorem for groups.

Let A, B and C be the given groups. To abbreviate we shall write

$$R = [(B, C), (C, A), (A, B)],$$

$$\bar{R} = ([B, C], [C, A], [A, B]),$$

and

$$T_A = [(A, [B, C]), (B, C)] = ([A, (B, C)], [B, C]),$$

while T_B and T_C are obtained by permutation of letters. We can now prove for any Dedekind structure

$$\bar{R} = [T_A, T_B] = [T_B, T_C] = [T_C, T_A],$$

$$R = (T_A, T_B) = (T_B, T_C) = (T_C, T_A).$$

This means that if we put

$$\mathfrak{A} = T_A/R, \qquad \mathfrak{B} = T_B/R, \qquad \mathfrak{C} = T_C/R,$$

the conditions (12) in Theorem 2 are satisfied, and hence we have the following interesting

THEOREM 4. *For any three normal subgroups A, B and C the quotient group*

$$\mathfrak{A} = ([A, B], [B, C], [C, A])/[(A, B), (B, C), (C, A)]$$

is Abelian.

A study of all possible similarity relations for quotients formed by consecutive elements in the structure generated by A, B and C reveals that these quotient groups fall into seven different classes of similarity. We shall state the results briefly. They may all be proved by application of the law of isomorphism.

[13] A. Speiser, *Theorie der Gruppen von endlicher Ordnung*, 2 Auflage, Berlin, 1927, p. 136.

THEOREM 5. *For any three normal subgroups A, B and C there exist the similarity relations*

$$(16) \qquad [A, B, C]/[A, B] \cong [B, C]/([A, B], [A, C]) \cong [C, (A, B)]/T_C \cong C/(C, [A, B]),$$

and two others obtained by permutation. Three dual sets of similarities are obtained by interchanging cross-cuts and unions in (16). *Finally, there exists the self-dual set of similarities*

$$(17) \qquad \begin{aligned} &([A, B], [A, C])/[A, (B, C)] \cong ([B, A], [B, C])/[B, (A, C)] \\ &\qquad \cong ([C, A], [C, B])/[C, (A, B)] \\ &\cong (A, [B, C])/[(A, B), (A, C)] \cong (B, [A, C])/[(B, A), (B, C)] \\ &\qquad \cong (C, [A, B])/[(C, A), (C, B)] \\ &\cong \bar{R}/T_A \cong \bar{R}/T_B \cong \bar{R}/T_C \\ &\cong T_A/R \cong T_B/R \cong T_C/R, \end{aligned}$$

and all the quotient groups (17) *are Abelian.*

The last remark follows from the proof of Theorem 4.

The similarity relations of Theorem 5 actually go back to Dedekind.[14] They have been restated in part by Remak[15] and Garrett Birkhoff.[16]

Another finite substructure of a general Dedekind structure may be defined by any four normal subgroups $A > a$ and $B > b$. In general the corresponding structure contains 18 elements and the consecutive quotients fall into 7 similar classes. Since most of these similarity relations are trivial, we shall only mention

THEOREM 6. *For any four normal subgroups $A > a$ and $B > b$ there exist the self-dualistic relations*

$$[a, b]/[a, (b, A)] \cong [b, (A, B)]/(A, B)$$

and

$$([a, B], [b, A])/[a, b] \cong (A, B)/[(a, B), (b, A)] \cong [a, (A, B)]/[a, (b, A)] \cong [b, (A, B)]/[b, (a, B)].$$

The last relations contain the lemma of Zassenhaus for normal subgroups.

It may be observed finally that the preceding relations can in some cases be extended to quasi-normal or permutable groups by using structure isomorphism or correspondences instead of similarity. In Chapter 2 we have already done this in the case of the lemma of Zassenhaus.

YALE UNIVERSITY.

[14] R. Dedekind, *Über die von drei Moduln erzeugte Dualgruppe*, Math. Ann., vol. 53 (1900); *Werke*, vol. 2, pp. 371–403.

[15] R. Remak, *Über Untergruppen direkter Produkte von drei Faktoren*, Journ. für Math., vol. 166 (1932), pp. 65–100.

[16] G. Birkhoff, *On the combination of subalgebras*, Proceedings Cambridge Phil. Soc., vol. 29 (1933), pp. 441–464.

A CLASS OF LINEAR GROUPS WITH INTEGRAL COEFFICIENTS

By Arthur B. Coble

Introduction. In the study of certain groups of Cremona transformations in S_k which I have called "regular groups" ([1]II, §§4, 5), the determination of the types of transformations in the Cremona group was accomplished by the use of a related linear group with integral coefficients. In later papers[2] it appeared that the class of Cremona groups with such related linear groups is probably quite extensive. It is the purpose of the present paper to discuss linear groups of this character independently of their association with Cremona groups. They are generated by a finite number of involutorial elements of a particular type, derived in §1, and characterized more completely in §2. These groups divide into two classes according to the values of a certain constant, $e = \pm 1$, and they have an additional integer parameter ϵ. The linear groups first mentioned occur when $\epsilon = 1$, $e = -1$. These occurred in pairs of "associated groups" and in §3 this isomorphism described as association is extended to generic ϵ.

Each group has an invariant linear, and an invariant quadratic, form. This imposes the necessary conditions on the coefficients of the generic element, which are obtained in §4. These conditions, however, are not sufficient. The de Jonquières subgroups, defined in §5 after the manner of the de Jonquières group of planar Cremona transformations, are used in §§7, 8 to separate the cases of finite and of infinite order.

The types of symmetric transformations, already determined when $\epsilon = 1$, $e = -1$, are found for the general case in §6.

A particular class, $g_\rho(\alpha)$, of these groups, whose generators are unusually simple, is studied in §9 with particular reference to aggregates of products of ternary de Jonquières transformations, and again in §10 with reference to the nature of the coefficients of its elements. The writer hopes to consider the more general group in a later paper.

1. **A particular type of involutorial matrix.** In connection with the so-called "symmetric Cremona transformations" there occur linear transformations with integer coefficients, the matrix of the coefficients having the particular form

$$
(1) \qquad \begin{array}{cccc}
\alpha & -\beta & -\beta & -\beta \cdots \\
\delta & -\gamma & -\epsilon & -\epsilon \cdots \\
\delta & -\epsilon & -\gamma & -\epsilon \cdots \\
\delta & -\epsilon & -\epsilon & -\gamma \cdots \\
\cdots & \cdots & \cdots & \cdots
\end{array}
$$

Received February 23, 1937.

[1] A. B. Coble, *Point sets and Cremona groups*, I: Trans. Amer. Math. Soc., vol. 16 (1915), pp. 155–198; II: Trans. Amer. Math. Soc., vol. 17 (1916), pp. 345–385.

[2] A. B. Coble, *Groups of Cremona transformations in space of planar type*, I: this Journal, vol. 2 (1936), pp. 1–9; II: this Journal, vol. 2 (1936), pp. 205–219.

where α, β, γ, δ, ϵ are positive integers and where γ may be zero. An example is the well-known transformation

$$(2)\qquad \begin{aligned} x_0' &= 2x_0 - x_1 - x_2 - x_3, \\ x_1' &= x_0 \phantom{{}- x_1} - x_2 - x_3, \\ x_2' &= x_0 - x_1 \phantom{{}- x_2} - x_3, \\ x_3' &= x_0 - x_1 - x_2, \end{aligned}$$

associated with the quadratic Cremona transformation in the plane.

In the instances mentioned these linear transformations are involutorial. We ask then under what conditions a square matrix of type (1) and order $r + 1$ is involutorial. The customary conditions yield the following:

$$(3)\qquad \begin{array}{ll} r = 1 & r > 1 \\ \alpha^2 - \beta\delta = 1, & \alpha^2 - r\beta\delta = 1, \\ \beta(-\alpha + \gamma) = 0, & -\alpha\beta + \beta\gamma + (r-1)\beta\epsilon = 0, \\ \delta(-\alpha + \gamma) = 0, & \alpha\delta - \gamma\delta - (r-1)\delta\epsilon = 0, \\ -\delta\beta + \gamma^2 = 1, & -\delta\beta + \gamma^2 + (r-1)\epsilon^2 = 1, \\ & -\delta\beta + 2\gamma\epsilon + (r-2)\epsilon^2 = 0. \end{array}$$

In the case $r > 1$, the last two conditions yield $(\gamma - \epsilon)^2 = 1$, whence the entire set can be written as

$$(4)\qquad \begin{array}{ll} \gamma = \epsilon + e & (e = \pm 1), \\ \alpha^2 = (r\epsilon + e)^2, & \delta[\alpha - (r\epsilon + e)] = 0, \\ \delta\beta = \epsilon(r\epsilon + 2e), & \beta[\alpha - (r\epsilon + e)] = 0. \end{array}$$

Since δ and β are to be greater than zero, we have

(5) *The matrix* (1), *if of two rows* $(r = 1)$, *is involutorial if* $\gamma = \alpha$ *and* δ, β *are complementary factors of* $\alpha^2 - 1 (\alpha > 1)$; *if of* $(r + 1)$ *rows* $(r > 1)$, *is involutorial if* $\gamma = \epsilon + e$ $(e = \pm 1)$, $\alpha = r\epsilon + e$, *and* δ, β *are complementary factors of* $\epsilon(r\epsilon + 2e)$, *the case* $r = 2$, $\epsilon = 1$, $e = -1$ *being excluded. We assume also that* $\epsilon \geqq 1$.

Certain particular cases of matrices (1) which are involutorial are excluded by the restrictions imposed in (5). These would be of no interest in connection with the groups about to be defined. We denote by $I_{12 \cdots r}$ the linear transformation on variables $x_0, x_1, \cdots, x_r$ whose matrix of coefficients is (1) as limited in (5).

It may be verified without difficulty that $I_{12 \cdots r}$ has the absolute invariants

$$(6)\qquad \begin{aligned} Q &= \delta x_0^2 - \beta(x_1^2 + \cdots + x_r^2), \\ L &= r\delta x_0 - (\alpha - 1)(x_1 + \cdots + x_r). \end{aligned}$$

It has the relatively invariant linear form

$$L' = r\delta x_0 - (\alpha + 1)(x_1 + \cdots + x_r), \tag{7}$$

with multiplier -1. It has also $r - 1$ linearly independent invariant forms of type $x_i - x_j$ $(i, j > 0; i \neq j)$, which are absolutely unaltered if $e = -1$, and reproduced with a multiplier -1 if $e = 1$. From these facts we conclude in geometric language that

(8) *The involution $I_{12\cdots r}$ with matrix* (1), *as limited in* (5), *is always of the type perspective at a point. Given the quadratic form Q, this point, or its polar linear form, defines the involution. If $r = 1$, this polar linear form is either L or L'; if $r > 1$, it is L or L' according as e is 1 or -1. If $r = 1$, the determinant of the matrix is -1; if $r > 1$, this determinant has the value $(-1)^r$ or -1 according as e is 1 or -1.*

2. **The linear groups** $G_\rho(\alpha)$, $G_\rho(r, \epsilon, e)$. Let $j_1, \cdots, j_r$ be any combination of r indices from $1, 2, \cdots, \rho$ $(\rho > r)$, and let $I_{j_1 \cdots j_r}$ be the involutorial linear transformation on variables $x_0, x_{j_1}, \cdots, x_{j_r}$ with the same matrix of coefficients as was assigned to $I_{1\cdots r}$ in **1**, the remaining variables in the set $x_1, \cdots, x_r$ being unaltered. We define the group $g_\rho(r, \epsilon, e)$ to be that generated by the $\binom{\rho}{r}$ involutions $I_{j_1 \cdots j_r}$ formed for the given $\epsilon \geqq 1$ and $e = \pm 1$. Thus $g_\rho(r, \epsilon, e)$ is a group of linear homogeneous transformations on $\rho + 1$ variables.

There is no loss of generality in setting

$$\delta = 1, \qquad \beta = \epsilon(r\epsilon + 2e) \tag{1}$$

in the generators $I_{j_1 \cdots j_r}$. For if $\delta \neq 1$, the transformation

$$x_0 = x_0', \qquad x_i = \delta x_i' \qquad (i = 1, \cdots, \rho) \tag{2}$$

carries the group into one for which $\delta = 1$. We have then

(3) *If $r = 1$, the linear group $g_\rho(\alpha)$ is generated by ρ involutions of the form*

$$I_j: \quad x_0' = \alpha x_0 - (\alpha^2 - 1)x_j, \qquad x_j' = x_0 - \alpha x_j, \qquad x_k' = x_k \quad (k \neq 0, j).$$

It has the invariant forms

$$Q = x_0^2 - (\alpha^2 - 1)(x_1^2 + \cdots + x_\rho^2), \qquad L = x_0 - (\alpha - 1)(x_1 + \cdots + x_\rho), \qquad (\alpha \geqq 2).$$

(4) *If $r > 1$, the linear group $g_\rho(r, \epsilon, e)$ on variables $x_0, x_1, \cdots, x_\rho$ is generated by $\binom{\rho}{r}$ involutions of the form*

$$I_{j_1\cdots j_r}: \quad \begin{aligned} x_0' &= (r\epsilon + e)x_0 - \epsilon(r\epsilon + 2e)(x_{j_1} + \cdots + x_{j_r}), \\ x_{j_i}' &= x_0 - \epsilon(x_{j_1} + \cdots + x_{j_r}) - e x_{j_i} \qquad (i = 1, \cdots, r), \\ x_k' &= x_k \qquad (k \neq 0,\ j_i). \end{aligned}$$

It has the invariant forms

$$Q = x_0^2 - \epsilon(r\epsilon + 2e)(x_1^2 + \cdots + x_\rho^2),$$

$$L = rx_0 - (r\epsilon + e - 1)(x_1 + \cdots + x_\rho).$$

Naturally, if $e = 1$, L may be taken in the simplified form

$$(5) \qquad L = x_0 - \epsilon(x_1 + \cdots + x_\rho) \qquad (e = 1).$$

Thus, for $r = 1$, there exists a $g_\rho(\alpha)$ for every $\alpha \geqq 2$ and for every $\rho \geqq 2$. For $r > 1$, there exists a $g_\rho(r, \epsilon, e)$ for every $\epsilon \geqq 1$, for every $\rho \geqq r + 1$, and for each $e = \pm 1$, except when $r = 2$, $\epsilon = 1$, $e = -1$.

It is clear from the nature of the generators that

(6) *If $r < \rho' < \rho$, the $g_{\rho'}(r)$ is a subgroup of $g_\rho(r)$, the other parameters α, or ϵ, e in the two groups being the same.*

We shall later (cf. **5** (1)) find other cases in which certain groups $g_{\rho'}$ are subgroups of another group, g_ρ.

If we consider that the generating involutions are harmonic perspectivities determined by skew S_k and $S_{\rho-1-k}$ in the S_ρ of $x_0, x_1, \cdots, x_\rho$, we see that

(7) *The group $g_\rho(\alpha)$ in S_ρ is generated by harmonic perspectivities whose spaces of fixed points are the S_0 with equation $(\alpha - 1)\xi_0 - \xi_i = 0$, and the $S_{\rho-1}$ with equation $x_0 - (\alpha + 1)x_i = 0$. If $e = -1$, the generators of $g_\rho(r, \epsilon, -1)$ have also, for spaces of fixed points, the S_0, $(r\epsilon - 2)\xi_0 - (\xi_{j_1} + \cdots + \xi_{j_r}) = 0$, and the $S_{\rho-1}$, $x_0 - \epsilon(x_{j_1} + \cdots + x_{j_r}) = 0$. If however $e = 1$, the generators of g_ρ have for spaces of fixed points the S_{r-1}, $L = x_{k_1} = x_{k_2} = \cdots = x_{k_{\rho-r}} = 0$, and the $S_{\rho-r}$, $L' = x_{j_1} - x_{j_2} = \cdots = x_{j_1} - x_{j_r} = 0$, where the indices $k_1, \cdots, k_{\rho-r}$ are complementary to $j_1, \cdots, j_r$ in $1, 2, \cdots, \rho$.*

3. **Associated groups.** In connection with regular Cremona groups and their attached linear groups $g_\rho(r, 1, -1)$ [cf. [1]II, §5 (27)], the author has called attention to an isomorphism referred to as "associated groups". This isomorphism persists in some measure for the more general linear groups considered here. We prove first that

(1) *If $\rho - r > 1$, the group $g_\rho(r, \epsilon, -1)$ is the linear transform of the group $g_\rho(\rho - r, \epsilon, -1)$, the generating involution $I_{j_1 \cdots j_r}$ of the first group being the transform of the generating involution $I_{j_{r+1} \cdots j_\rho}$ of the second group.*

For the proof it is necessary only to find a linear transformation T of non-zero determinant from x to $\bar{x}$ which converts the fixed S_0, $S_{\rho-1}$ of $I_{j_1 \cdots j_r}$ (cf. **2** (7)), i.e.,

$$(r\epsilon - 2)\xi_0 - (\xi_{j_1} + \cdots + \xi_{j_r}), \qquad x_0 - \epsilon(x_{j_1} + \cdots + x_{j_r}),$$

into the fixed S_0', $S_{\rho-1}'$ of $I_{j_{r+1} \cdots j_\rho}$, which are

$$\{(\rho - r)\bar{\epsilon} - 2\}\bar{\xi}_0 - (\bar{\xi}_{j_{r+1}} + \cdots + \bar{\xi}_{j_\rho}), \qquad \bar{x}_0 - \bar{\epsilon}(\bar{x}_{j_{r+1}} + \cdots + \bar{x}_{j_\rho}).$$

Since T must have the same effect for all sets of complementary indices $j_1, \cdots, j_r$ and $j_{r+1}, \cdots, j_\rho$, T must be symmetric in $x_1, \cdots, x_\rho$, and we set

$$(2) \qquad T: \quad \begin{aligned} \bar{x}_0 &= lx_0 + m(x_1 + \cdots + x_\rho), \\ \bar{x}_i &= nx_0 + p(x_1 + \cdots + x_\rho) + qx_i \qquad (i = 1, \cdots, \rho). \end{aligned}$$

The determinant of T is $q^{\rho-1}[lq + \rho lp - \rho mn]$. The conditions that S_0, S_0' and $S_{\rho-1}$, $S_{\rho-1}'$ are conjugate respectively under T and T^{-1} respectively are

$$l(r\epsilon - 2) + mr:n(r\epsilon - 2) + pr + q:n(r\epsilon - 2) + pr = (\rho - r)\bar{\epsilon} - 2:0:\bar{\epsilon},$$

$$l - \bar{\epsilon}(\rho - r)n:m - \bar{\epsilon}(\rho - r)p:m - \bar{\epsilon}(\rho - r)p - \bar{\epsilon}q = 1:-\epsilon:0.$$

These four conditions on l, m, n, p, q respectively have the matrix

$$\begin{array}{ccccc} 0 & 0 & r\epsilon - 2 & r & 1 \\ 0 & -1 & 0 & \bar{\epsilon}(\rho - r) & \bar{\epsilon} \\ r\epsilon - 2 & r & 0 & 0 & (\rho - r)\bar{\epsilon} - 2 \\ \epsilon & 0 & -\epsilon\bar{\epsilon}(\rho - r) & 0 & \bar{\epsilon}. \end{array}$$

If l and m are eliminated from the last three, and the result simplified by using the first condition, we get $(\bar{\epsilon} - \epsilon)q = 0$. Since q, a factor of the determinant of T, is not zero, $\bar{\epsilon} = \epsilon$. The conditions are then dependent. Setting $p = 0$, we find that

(3) *The groups* $g_\rho(r, \epsilon, -1)$, $g_\rho(\rho - r, \epsilon, -1)$ *of* (1) *are conjugate under* T *in* (2) *with* $l:m:n:p:q = -(\rho\epsilon - 2):\epsilon(r\epsilon - 2):-1:0:(r\epsilon - 2)$, *and with determinant* $2(r\epsilon - 2)^\rho > 0$.

If $\rho - r = 1$, we seek a transformation T which converts the fixed S_0, S_r of $I_{j_1 \cdots j_r}$ as given above into the fixed S_0', S_r' of $I_{j_{r+1}}$ of $g_{r+1}(\alpha)$, which are $(\alpha - 1)\bar{\xi}_0 + \bar{\xi}_{j_{r+1}} = 0$, $\bar{x}_0 - (\alpha + 1)\bar{x}_{j_{r+1}}$. The matrix of the four conditions on l, m, n, p, q is

$$\begin{array}{ccccc} 0 & -1 & 0 & \alpha + 1 & \alpha + 1 \\ \epsilon & 0 & -\epsilon(\alpha + 1) & 0 & \alpha + 1 \\ 0 & 0 & r\epsilon - 2 & r & 1 \\ r\epsilon - 2 & r & 0 & 0 & \alpha - 1. \end{array}$$

If the first and second equations are solved for l, m, and the results substituted in the last, a comparison with the third equation yields $[(\alpha + 1) - \epsilon]q = 0$, whence $\epsilon = \alpha + 1$. Again setting $p = 0$, we find that

(4) *For* $\epsilon \geqq 3$, *the group* $g_{r+1}(r, \epsilon, -1)$ *can be transformed into the group* $g_{r+1}(\alpha = \epsilon - 1)$ *by the transformation* T *in* (2) *with* $l:m:n:p:q = -(r\epsilon - 2) - \epsilon:\epsilon(r\epsilon - 2):-1:0:r\epsilon - 2$, *and with determinant* $2(r\epsilon - 2)^{r+1} > 0$ *in such wise that generators* $I_{j_1 \cdots j_r}$ *and* $I_{j_{r+1}}$ *of the respective groups are conjugate.*

It is clear that no group $g_\rho(r, \epsilon, e)$ with $e = 1$ can be transformed into a group $g_\rho(\alpha)$ or into a $g_\rho(r, \epsilon, e)$ with $e = -1$ in such a way that the generating involutions of the one group pass into those of the other, since, according to **2** (7), the fixed spaces of the generators of the one group are of different character

from those of the generators of the other. Nor, for the same reason, can two g_ρ's with $e = 1$ be transformed into each other in such fashion that complementary generators pass into each other. Thus *association* is characteristic of the groups $g_\rho(\alpha)$ and $g_\rho(r, \epsilon, -1)$.

(5) *The group $g_\rho(r, \epsilon, -1)$ will have not only the series of subgroups isomorphic with $g_{\rho-k}(r, \epsilon, -1)$ $(k = 1, \cdots, \rho - r - 1)$ mentioned in* **2** (6) *but also subgroups isomorphic with $g_{\rho-k-l}(\rho - r - k, \epsilon, -1)$ $(l = 1, \cdots, r - 1)$. If however $\epsilon = 1, 2$, then $k = 1, \cdots, \rho - r - 2$.*

For, according to **2** (6), $g_\rho(r, \epsilon, -1)$ contains subgroups $g_{\rho-k}(r, \epsilon, -1)$ obtained by fixing k of the variables. The subgroup $g_{\rho-k}(r, \epsilon, -1)$ is, according to (3), isomorphic with $g_{\rho-k}(\rho - r - k, \epsilon, -1)$, and this latter, with l variables fixed, yields subgroups $g_{\rho-k-l}(\rho - r - k, \epsilon, -1)$.

The distinguishing feature of the association here derived is the correspondence between complementary generators of the associated groups which automatically sets up an isomorphism. There may well be other types of isomorphism between two of these groups, and even other cases of conjugacy under linear transformation.

Reverting to (3), we observe that if $\rho = 2r$, the associated groups coincide, whence

(6) *The groups $g_{2r}(r, \epsilon, -1)$ are self-associated. Each possesses an inner isomorphism in which complementary generators correspond.*

4. **Relations connecting the integer coefficients of the elements of the groups g_ρ.** In order to include both the groups $g_\rho(\alpha)$ and the groups $g_\rho(r, \epsilon, e)$, it is convenient to return to the notation of **1** with

$$\gamma = \epsilon + e, \qquad \delta\beta = \epsilon(r\epsilon + 2e), \qquad r\delta\beta = \alpha^2 - 1.$$

We also take the invariant forms Q, L as in **1** (6).

Let the generic element of g_ρ be written as

$$(1) \qquad T: \begin{aligned} x_0' &= \alpha_{00}x_0 - \alpha_{01}x_1 \cdots - \alpha_{0j}x_j \cdots - \alpha_{0\rho}x_\rho, \\ x_i' &= \alpha_{i0}x_0 - \alpha_{i1}x_1 \cdots - \alpha_{ij}x_j \cdots - \alpha_{i\rho}x_\rho \qquad (i = 1, \cdots, \rho). \end{aligned}$$

We first observe that

$$(2) \qquad \alpha_{00}^2 = 1 + a_0\delta\beta, \qquad \alpha_{i0} = \delta d_i, \qquad \alpha_{0j} = \beta b_j \qquad (i > 0, j > 0).$$

For these relations are satisfied by the generators, and, if satisfied by (1), are also satisfied by the product of (1) and a generator. Hence they are true of every element.

The quadratic relations on the coefficients α_{ij} are as follows:

$$(i, j, k = 1, \cdots, \rho)$$

$$\Sigma_i d_i^2 = a_0, \qquad \Sigma_j b_j^2 = a_0,$$

$$(3) \qquad \Sigma_i \alpha_{ij}^2 = \delta\beta b_{ij}^2 + 1, \qquad \Sigma_j \alpha_{ij}^2 = \delta\beta d_i^2 + 1,$$

$$\Sigma_i d_i \alpha_{ij} = b_j \cdot \alpha_{00}, \qquad \Sigma_j b_j \alpha_{ij} = d_i \cdot \alpha_{00},$$

$$\Sigma_i \alpha_{ij} \alpha_{ik} = \delta\beta b_j b_k \ (j \neq k), \qquad \Sigma_j \alpha_{ij} \alpha_{kj} = \delta\beta d_i d_k \ (i \neq k).$$

The first column of four relations expresses that Q is unaltered by T. But it also expresses that the matrix T' obtained by interchanging b_i with d_i, also α_{ij} with α_{ji}, in T satisfies the identity $TT' = 1$, whence T' is the inverse of T. Hence

(4) *The inverse of T is obtained from T by interchanging b_i with d_i, and α_{ij} with α_{ji}.*

The second column of four relations (3) then expresses that Q is unaltered by T^{-1}.

If we apply T and T^{-1} to the invariant form L, the following four linear relations on the coefficients of T are obtained:

$$(5) \qquad \begin{aligned} (\alpha - 1)\Sigma_i d_i &= r(\alpha_{00} - 1), \qquad & (\alpha - 1)\Sigma_j b_j &= r(\alpha_{00} - 1), \\ \Sigma_i \alpha_{ij} &= (\alpha + 1)b_j - 1, \qquad & \Sigma_j \alpha_{ij} &= (\alpha + 1)d_i - 1. \end{aligned}$$

The following fact is needed in the next section:

(6) *If T in $g_\rho(r, \epsilon, e)$ is a product of g generators, then*

$$\alpha_{00} = (-e)^g + \mu(r\epsilon + 2e).$$

For this relation is true for a single generator, $g = 1$, $\mu = 1$, $\alpha_{00} = \alpha = r\epsilon + e$. Let us assume that it is true for T, i.e., that $\alpha_{00} \equiv (-e)^g \bmod (r\epsilon + 2e)$. It is then true for the product T' of T and an additional generator, since $\alpha'_{00} = \alpha_{00}(r\epsilon + e) - \epsilon(r\epsilon + 2e)(d_{i_1} + \cdots + d_{i_r})$ and $\alpha'_{00} \equiv \alpha_{00}(r\epsilon + e) \equiv \alpha_{00}(-e) \equiv (-e^g)(-e) \equiv (-e)^{g+1} \bmod (r\epsilon + 2e)$.

The corresponding theorem for $g_\rho(\alpha)$ is

(7) *If T in $g_\rho(\alpha)$ is a product of g generators with determinant $\Delta = (-1)^g$, then $\alpha_{00} = \Delta + \mu(\alpha + 1)$.*

This indeed is true of the generators themselves, and it remains true when they are combined.

A further property, used in **9**, of T in $g_\rho(\alpha)$ is

(8) *The value μ defined in* (7) *satisfies the relation*

$$\mu - \Sigma_j b_j = 2\lambda \qquad (\lambda \text{ integral and } \geqq 0).$$

To prove this we consider the product TI_i with coefficients $\alpha'_{00} = (n - 2)\alpha_{00} - (n - 1)(n - 3)d_i$, $b'_j = (n - 2)b_j - \alpha_{ij}$. Since, for the product, $\Delta' = -\Delta$, $\alpha'_{00} = (n - 2)\Delta + (n - 2)\mu(n - 1) - (n - 1)(n - 3)d_i = -\Delta + (n - 1)\{\Delta + \mu(n - 2) - (n - 3)d_i\}$, whence $\mu' = \Delta + \mu(n - 2) - (n - 3)d_i$. Also, according to (5), $\Sigma_j b'_j = (n - 2)\Sigma_j b_j - (n - 1)d_i + 1$. Hence $\mu' - \Sigma_j b'_j = (n - 2)(\mu - \Sigma_j b_j) + \Delta - 1 + 2d_i$. Since $\Delta - 1$ is even, $\mu' - \Sigma_j b'_j$ is even if $\mu - \Sigma_j b_j$ is even. But, for a single generator, $\mu = 1$ and $\Sigma_j b_j = 1$, whence $\mu - \Sigma_j b_j$ is even, and remains even under multiplication of the generators.

5. **De Jonquières subgroups of** $g_\rho(r, \epsilon, e)$. We have found in **3** that when $e = -1$, the groups $g_\rho(r, \epsilon, -1)$ and $g_\rho(\rho - r, \epsilon, -1)$ are simply isomorphic in such fashion that, if $i_1 \cdots i_r$ and $j_1 \cdots j_{\rho-r}$ are complementary sets of indices from $1, \cdots, \rho$, the generators $I_{i_1 \cdots i_r}$ and $I_{j_1 \cdots j_{\rho-r}}$ of the two groups correspond. If we take in $g_\rho(\rho - r, \epsilon, -1)$ that subgroup mentioned in **2** (6) which leaves $x_1, \cdots, x_k$ unaltered, the isomorphic subgroup in $g_\rho(r, \epsilon, -1)$ is generated by involutions $I_{1 \cdots k i_1 \cdots i_{r-k}}$, where the indices $i_1 \cdots i_{r-k}$ are selected from $k + 1, \cdots, \rho$. But this set of generators will generate a subgroup whether e be 1 or -1. We call this type of subgroup a de Jonquières subgroup because of its first appearance in connection with de Jonquières planar transformations.
(1) *The de Jonquières subgroup of* $g_\rho(r, \epsilon, e)$ *generated by involutions the* $I_{12 \cdots k i_1 \cdots i_{r-k}}$ *is simply isomorphic with* $g_{\rho-k}(r - k, \epsilon, e)$ $(k < r)$.

It is clear that, due to the relations $x'_m - x'_n = (-e)^g(x_m - x_n)$ $[m, n = 1, \cdots, k;\ m < n]$, which persist under repeated application (g times) of the given generators, the elements of the subgroup will have like rows, $x'_1, \cdots, x'_k$, except in the rectangle of coefficients under $x_1, \cdots, x_k$. In this rectangle of coefficients the principal diagonal terms will have the additive part $(-e)^g$. It will thus be sufficient to have the sum of these rows and we introduce the variable

$$\sigma = x_1 + \cdots + x_k. \tag{2}$$

A typical generator then reads

$$(3)\qquad I_{1 \cdots k, k+1, \cdots, r}:\ \begin{aligned} x'_0 &= (r\epsilon + e)x_0 - \epsilon(r\epsilon + 2e)\sigma - \epsilon(r\epsilon + 2e) \\ &\qquad\qquad (x_{k+1} + \cdots + x_r), \\ \sigma' &= kx_0 - (k\epsilon + e)\sigma - k\epsilon(x_{k+1} + \cdots + x_r), \\ x'_j &= x_0 - \epsilon\sigma - \epsilon(x_{k+1} + \cdots + x_r) - ex_j \\ &\qquad\qquad (j = k + 1, \cdots, r). \end{aligned}$$

This transformation on the $r - k + 2$ variables x_0, σ, x_j is itself involutorial.

With generators of the type (3) we make the following change of variables:

$$(4)\qquad \begin{aligned} t &= kx_0 - (r\epsilon + 2e)\sigma, \qquad z_j = x_j \qquad (j = k + 1, \cdots, \rho). \\ z_0 &= x_0 - \epsilon\sigma. \end{aligned}$$

The transformation inverse to (4) is

$$(5)\qquad \begin{aligned} \{(r - k)\epsilon + 2e\}x_0 &= (r\epsilon + 2e)z_0 - \epsilon t, \qquad x_j = z_j, \\ \{(r - k)\epsilon + 2e\}\sigma &= kz_0 - t. \end{aligned}$$

In these new variables the equation of the generator (3) is

$$(6)\qquad \begin{aligned} t' &= (-e)t, \quad z'_l = z_l \qquad\qquad (l = r + 1, \cdots, \rho), \\ z'_0 &= \{(r - k)\epsilon + e\}z_0 - \epsilon\{(r - k)\epsilon + 2e\}(z_{k+1} + \cdots + z_r), \\ z'_j &= z_0 - \epsilon(z_{k+1} + \cdots + z_r) - ez_j \qquad (j = k + 1, \cdots, r). \end{aligned}$$

Thus to the generator (3) of the de Jonquières subgroup there corresponds a generator $I_{k+1,\cdots,r}$ of the $g_{\rho-k}(r-k,\epsilon,e)$, and thereby to every element of the one group there corresponds an element of the other.

It is of interest to have the expression for the element of $g_\rho(r,\epsilon,e)$ which corresponds to a given element of $g_{\rho-k}(r-k,\epsilon,e)$. Let then the generic element of $g_{\rho-k}(r-k,\epsilon,e)$ be

$$(7)\qquad \begin{aligned} z_0' &= \beta_{00}z_0 - \Sigma_j\beta_{0j}z_j,\\ z_i' &= \beta_{i0}z_0 - \Sigma_j\beta_{ij}z_j \qquad (i,j=k+1,\cdots,\rho),\end{aligned}$$

and let it be generated by g generators. According to **4** (2), (6)

$$(8)\qquad \beta_{00} = (-e)^g + \mu'\{(r-k)\epsilon + 2e\},\qquad \beta_{0j} = \{(r-k)\epsilon+2e\}\epsilon b_j'.$$

We have from (5) that $\{(r-k)\epsilon + 2e\}x_0' = (r\epsilon+2e)z_0' - \epsilon t'$. On replacing z_0' in this from (7), and t' by $(-e)^g t$ [cf. (6)], then making use of (4) and applying the relations (8), we can factor out $\{(r-k)\epsilon+2e\}$ to obtain the value of x_0' in (9). The value of σ', similarly obtained, is

$$\sigma' = k\mu'x_0 - [\epsilon k\mu' - (-e)^g]\sigma - k\epsilon\Sigma_j b_j'x_j \qquad (j=k+1,\cdots,\rho).$$

In the light of our introductory remarks this yields the value of x_i' in (9). The remaining formula in

$$(9)\qquad \begin{aligned} x_0' &= [(-e)^g + \mu'(r\epsilon+2e)]x_0 - \epsilon(r\epsilon+2e)\mu'(x_1+\cdots+x_k)\\ &\qquad\qquad - \epsilon\,(r\epsilon+2e)\Sigma_j b_j'x_j,\\ x_i' &= \mu'x_0 - \epsilon\mu'(x_1+\cdots+x_k) + (-e)^g x_i - \epsilon\Sigma_j b_j'\,x_j,\\ x_l' &= \beta_{l0}x_0 - \epsilon\beta_{l0}(x_1+\cdots+x_k) - \Sigma_j\beta_{lj}x_j\\ &\qquad\qquad (i=1,\cdots,k;\ j,l=k+1,\cdots,\rho)\end{aligned}$$

is almost immediate. Hence

(10) *If a generic element of* $g_{\rho-k}(r-k,\epsilon,e)$ *is given as in* (7), (8), *then* (9) *is a generic element of a de Jonquières subgroup of a* $g_\rho(r,\epsilon,e)$.

In another connection I have called this process of passing from a given group to a subgroup of a more comprehensive group, the "dilation" of the given group [cf. [3]§4].

In the above we have set the limit $k<r$. If however $k=r-1$, the $g_{\rho-k}(r-k,\epsilon,e)$ would be a $g_{\rho-k}(\alpha)$ with only one free variable in the generators. This case would however be included under the above argument by setting

$$(11)\qquad \begin{gathered} g_{\rho-k}(1,\epsilon,e) = g_{\rho-k}(\alpha=\epsilon+e),\\ \epsilon \geqq 3 \text{ if } e=-1,\qquad \epsilon\geqq 1 \text{ if } e=1.\end{gathered}$$

[3] A. B. Coble, *The ten nodes of the rational sextic and the Cayley symmetroid*, Amer. Jour., vol. 41 (1919), pp. 243–265.

6. **Symmetric transformations and related theorems.** It may happen that the groups $g_\rho(\alpha)$, $g_\rho(r, \epsilon, e)$ contain other elements of the type **1** (1) than the generators. Such elements are called "symmetric transformations". Thus the $g_\rho(3, 1, -1)$ generated by elements of type **1** (2), whose elements represent the effect of Cremona transformations in the plane, contains three such symmetric transformations corresponding respectively to Cremona transformations of orders 5, 8, 17 with 6, 7, 8 F-points of orders 2, 3, 6. Such symmetric transformations in a given group will generate a $g_\rho(r', \epsilon', e')$ which is a subgroup of the given group and which therefore has the same invariants, Q, L.

We cannot assume that $\delta = 1$ in these symmetric transformations, and will therefore denote them more specifically by $I(r', \alpha', \delta', \beta', \epsilon', e')$, where $\alpha' = r'\epsilon' + e'$, and $\delta'\beta' = \epsilon'(r'\epsilon' + 2e')$. Consider first the Q, L of $g_\rho(\alpha)$ in **2** (3). They will be unaltered by this I if

$$(\alpha' + 1) = \delta'(\alpha + 1), \qquad \beta' = (\alpha^2 - 1)\delta', \qquad \alpha' - 1 = r'\delta'(\alpha - 1).$$

From the first and last we see that δ' is a divisor of 2. If $\delta' = 1$, $\alpha' = \alpha$, $r' = 1$, and $\beta' = \alpha^2 - 1$. This yields merely a generator I_1 of $g_\rho(\alpha)$. If $\delta' = 2$, $\alpha' = 2\alpha + 1$, and $\alpha = r'(\alpha - 1)$, whence $\alpha = r' = 2$, $\beta' = 6$. From $\epsilon'(r'\epsilon' + 2e') = \beta'\delta' = 12$, and $\alpha' = r'\epsilon' + e' = 5$, we get $\epsilon'(5 + e') = 12$, $2\epsilon' + e' = 5$. Thus, for $e' = \pm 1$, we have two solutions, $e' = -1$, $\epsilon' = 3$, and $e' = 1$, $\epsilon' = 2$. These solutions yield two matrices of type **1** (1):

$$M: \begin{matrix} 5 & -6 & -6 \\ 2 & -2 & -3 \\ 2 & -3 & -2; \end{matrix} \qquad M': \begin{matrix} 5 & -6 & -6 \\ 2 & -3 & -2 \\ 2 & -2 & -3. \end{matrix} \tag{1}$$

If we examine the group g_2 ($\alpha = 2$) with generators I_1, I_2, we find that I_1I_2 has the period three, whence $g_2(2)$ is dihedral of order six. The remaining element of period two is $I_1I_2I_1 = I_2I_1I_2$ with matrix M. Naturally an element with matrix M' will not occur in the $g_2(2)$. Hence

(2) *The $g_\rho(2)$ is the only $g_\rho(\alpha)$ with a symmetric transformation. The $g_\rho(2, 3, -1)$ generated by this transformation with matrix M is a subgroup of $g_\rho(2)$. In particular $g_2(2)$ is dihedral of order six, and M is in the conjugate set of three involutions.*

If the $g_\rho(2, 3, -1)$ were amplified by introducing the permutation group of $x_1, \cdots, x_\rho$, the larger group so obtained would contain the $g_\rho(2, 2, 1)$ generated by involutions M', since M' is M followed or preceded by the interchange of x_1, x_2. We have not of course proved that $g_\rho(2, 3, -1)$ does not contain $\Pi_{\rho!}$. This however seems to be unlikely. We prove later that, when $\epsilon = 1$, either $\Pi_{\rho!}$ or its even subgroup is contained in $g_\rho(r, 1, e)$.

Suppose now that $r > 1$, and let $g_\rho(r, \epsilon, e)$ be generated as in **2** (4). We ask for generating involutions with $r' \geqq r$, ϵ', e', and δ', $\beta' = \epsilon'(r'\epsilon' + 2e')$, which leave the Q, L in **2** (4) unaltered. In order to bring in α, α' we observe that, since $e^2 = e'^2 = 1$, and $\alpha = r\epsilon + e$, $\alpha' = r'\epsilon' + e'$,

$$r\epsilon(r\epsilon + 2e) = r\beta = \alpha^2 - 1,$$
$$r'\epsilon'(r'\epsilon' + 2e') = r'\delta'\beta' = \alpha'^2 - 1.$$

It is also convenient to write Q, L as

$$rQ = rx_0^2 - (\alpha^2 - 1)(x_1^2 + \cdots + x_\rho^2),$$
$$L = rx_0 - (\alpha - 1)(x_1 + \cdots + x_\rho).$$

Then the conditions for invariance ultimately reduce to the following three:

$$(3) \qquad \begin{gathered} (r' - r)(\alpha'\alpha - 1) = (\alpha' - \alpha)(r' + r), \\ \delta' = (\alpha' + 1)/(\alpha + 1), \qquad r\beta' = (\alpha' + 1)(\alpha - 1). \end{gathered}$$

Since $r' \geqq r$, we first consider $r' = r$. Then from (3), $\alpha' = \alpha$, $\delta' = 1$, $\beta' = \beta$. Also, from $\alpha' = \alpha$, we have $r(\epsilon' - \epsilon) = e - e'$. The obvious case is $\epsilon' = \epsilon$, $e' = e$, which yields no new generator. But it may be that $e' = -e$, or $r(\epsilon' - \epsilon) = 2e$. Since $e = \pm 1$, and $r \geqq 2$, then $r = 2$, $\epsilon' = \epsilon + e$, $\epsilon' + e' = \epsilon$. Thus the new involution is merely the generator I_{12} with the last two rows (or the last two columns) interchanged as M, M' above. As noted above, $g'_\rho(2, \epsilon, -e)$ generated by involutions I'_{12} is not the same as the $g_\rho(2, \epsilon, e)$ generated by involutions I_{12}. We also have excluded in **1** (5) the case $\epsilon = 1$, $e = -1$.

Suppose now that $r' > r$, say $r' = r + i$ $(i \geqq 1)$. Suppose also that $e = 1$. Then the first equation (6) reads

$$\alpha'(i\epsilon - 2) = -\epsilon(2r + i) - 2.$$

Hence $i\epsilon - 2 < 0$, i.e., $i = 1$, $e = 1$. Hence $\alpha = r + 1$, $r' = r + 1$, $\alpha' = 2r + 3$, $\delta' = 2$, $\beta' = 2r + 4$. From $\alpha' = 2r + 3 = r'\epsilon' + e' = (r + 1)\epsilon' + e'$, we have $e' = 1$, $\epsilon' = 2$, since $e' = -1$ would yield a fractional ϵ. This is the case (c) of Theorem (4).

Suppose that $r' > r$ as before, and that $e = -1$. Then the first equation (6) reads

$$\alpha'(ir\epsilon - 2r - 2i) = -2r(r\epsilon - 1) - i(r\epsilon - 2).$$

Since $r \geqq 2$, $\epsilon \geqq 1$, $i \geqq 1$, then $ir\epsilon - 2r - 2i < 0$. If t, t' are respectively the larger and smaller of r, i, then $\epsilon t' - 4 < 0$ or $\epsilon \geqq 3$. If $\epsilon = 3$, $t' = i = 1$, since $r \geqq 2$. But then $r - 2 < 0$. Again if $\epsilon = 2$, $t' = i = 1$. This yields a case, $\epsilon = 2$, $r' = r + 1$, $\alpha' = 2r^2 - 1$, $\alpha = 2r - 1$, $\delta' = r$, $\beta' = 4r(r - 1)$, listed as (d) in Theorem (4). From $\alpha' = r'\epsilon' + e'$, or $2r^2 - 1 = (r + 1)\epsilon' + e'$, we find that $e' = 1$, $\epsilon' = 2(r - 1)$.

There remain the cases for which $e = -1$ and $\epsilon = 1$. These however are the groups $g_\rho(r, \epsilon, e)$ attached to the groups of "regular Cremona transformations" [cf. [1]II, §4], and for these groups the symmetric transformations have already been determined [cf. [1]II, §5, pp. 368–9]. These are listed in Theorem (4) as (e), $\cdots$, (h).

(4) *The following sets of involutions* $I(r, \alpha, \delta, \beta, \epsilon, e)$ *have the same invariant* Q, L:

(a1): $(1, 2, 1, 3, -, -)$

(a2): $(2, 5, 2, 6, 3, -1)$

(a3): $(2, 5, 2, 6, 2, 1)$

(b1): $(2, 2\epsilon + e, 1, 2\epsilon(\epsilon + e), \epsilon, e)$

(b2): $(2, 2\epsilon + e, 1, 2\epsilon(\epsilon + e), \epsilon + e, -e)$

(c1): $(r, r + 1, 1, r + 2, 1, 1)$

(c2): $(r + 1, 2r + 3, 2, 2(r + 2), 2, 1)$

(d1): $(r, 2r - 1, 1, 4(r - 1), 2, -1)$

(d2): $(r + 1, 2r^2 - 1, r, 4r(r - 1), 2(r - 1), 1)$

(e1): $(3, 2, 1, 1, 1, -1)$

(e2): $(6, 5, 2, 2, 1, -1)$

(e3): $(7, 8, 3, 3, 1, 1)$

(e4): $(8, 17, 6, 6, 2, 1)$

(f1): $(4, 3, 1, 2, 1, -1)$

(f2): $(6, 7, 2, 4, 1, 1)$

(f3): $(7, 15, 4, 8, 2, 1)$

(g1): $(5, 4, 1, 3, 1, -1)$

(g2): $(8, 49, 10, 30, 6, 1)$

(h1): $(2k, 2k - 1, 1, 2(k - 1), 1, -1)$

(h2): $(2(k + 1), 2k^2 - 1, k, 2k(k - 1), k - 1, 1)$.

We have seen that the cases (a), (b) were not of special interest, and that cases (e), $\cdots$, (h) have been discussed. We consider then more particularly the cases (c), (d). In both we pass from $x_1, \cdots, x_r$ to $x_1, \cdots, x_{r+1}$. Let then J_k be the generating involution $I_{1,\cdots,k-1,k+1,\cdots,r+1}$ $(k = 1, \cdots, r + 1)$. With $\beta = \epsilon(r\epsilon + 2e)$, the product $J_{r+1}J_r$ has the matrix

$$(5) \quad J_{r+1}J_r: \quad \begin{array}{cccccc} 1 + \beta & -\beta\epsilon & -\beta\epsilon & \cdots & -\beta(\epsilon + e) & -\beta \\ \epsilon & -(\epsilon^2 - 1) & -\epsilon^2 & \cdots & -\epsilon(\epsilon + e) & -\epsilon \\ \epsilon & -\epsilon^2 & -(\epsilon^2 - 1) & \cdots & -\epsilon(\epsilon + e) & -\epsilon \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 1 & -\epsilon & -\epsilon & \cdots & -(\epsilon + e) & 0 \\ \epsilon + e & -\epsilon(\epsilon + e) & -\epsilon(\epsilon + e) & \cdots & -\epsilon(\epsilon + 2e) & -(\epsilon + e). \end{array}$$

If in this $\epsilon + e = 1$, since $\epsilon > 0$, $e = \pm 1$, then $\epsilon = 2$, $e = -1$, $\beta = 4(r - 1)$, as in case (d1). From an inspection of the last two rows and last two columns it is clear that then

$$(6) \qquad J_{r+1}J_r = J_rJ_{r+1}.$$

Hence

(7) *In the case* (d1) *of* (4), *the generating involutions are permutable. Their product* $J_1J_2 \cdots J_{r+1}$ *is a symmetric involution necessarily of type* (d2).

In connection with the cases (c) we use the element

$$
(8)\quad J_{r+1}J_rJ_{r+1}:\quad
\begin{array}{cccccc}
\alpha' & -\beta(\epsilon e+\eta) & -\beta(\epsilon e+\eta) & \cdots & -\beta(2\epsilon e-e+\eta) & -\beta(\epsilon+e) \\
\epsilon e+\eta & -(\epsilon^2 e+e+\epsilon\eta) & -\epsilon(\epsilon e+\eta) & \cdots & -(\epsilon+e)(\epsilon e+\eta-1) & -\epsilon(\epsilon+e) \\
\epsilon e+\eta & -\epsilon(\epsilon e+\eta) & -(\epsilon^2 e+e+\epsilon\eta) & \cdots & -(\epsilon+e)(\epsilon e+\eta-1) & -\epsilon(\epsilon+e) \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
2\epsilon e-e+\eta & -(2\epsilon^2 e-\epsilon e+\epsilon\eta) & -(2\epsilon^2 e-\epsilon e+\epsilon\eta) & \cdots & -(\epsilon+e)(2\epsilon e-e+\eta-1) & -\epsilon(\epsilon+2e) \\
\epsilon+e & -\epsilon(\epsilon+e) & -\epsilon(\epsilon+e) & \cdots & -\epsilon(\epsilon+2e) & -(\epsilon+e),
\end{array}
$$

where $\alpha' = r\epsilon + e + \beta(\epsilon + e - 1)$, $\eta = \epsilon^2 - \epsilon + 1$, $\beta = \epsilon(r\epsilon + 2e)$. This matrix will be of the typically involutorial form only if the elements to the right of α' are all equal. This requires

$$\epsilon + e = \epsilon e + \epsilon^2 - \epsilon + 1 = 2\epsilon e - e + \epsilon^2 - \epsilon + 1,$$

whence $\epsilon = 1$. Then also the elements below α' are equal. Moreover the last two elements in the second row must be equal, which requires further that $e = 1$. Then (8) is the involution $I_{1,\cdots,r+1}$ described in (c2) of (4), except that either the last two rows or the last two columns are interchanged. Hence

$$(9)\quad J_rJ_{r+1}J_r = J_{r+1}J_rJ_{r+1} = I_{1,\cdots,r+1}\cdot(r, r+1) = (r, r+1)\cdot I_{1,\cdots,r+1}.$$

From this there follows at once that

(10) *In the* $g_{r+1}(r, 1, 1)$, *the product* J_iJ_j $(i, j = 1, \cdots, r+1)$ *has the period three, and the transform* $J_iJ_jJ_i$ *the period two and matrix of type* (8). *This linear group contains the subgroup of even permutations of the variables* $x_1, \cdots, x_{r+1}$, *the cyclic elements,* J_iJ_j *and* (ijk), *of period three, being in the same conjugate set.*

For $J_rJ_{r+1}J_r = I_{1,\cdots,r+1}\cdot(r, r+1)$, $J_rJ_{r-1}J_r = (r, r-1)\cdot I_{1,\cdots,r+1}$, whence $J_r\cdot J_{r-1}J_{r+1}\cdot J_r = (r, r-1)\cdot(r, r+1) = (r-1, r+1, r)$.

An examination of the matrix (8) for the case $\epsilon = 1$, $e = -1$ shows that then $J_{r+1}J_rJ_{r+1} = (r, r+1)$. Hence

(11) *The group* $g_{r+1}(r, 1, -1)$ *contains as a subgroup the permutation group* $\Pi_{(r+1)!}$ *of* $x_1, \cdots, x_{r+1}$, *the transpositions* (x_ix_j) *being in the same conjugate set as the generators.*

Since $g_{\rho'}(r, \epsilon, e)$ is a subgroup of $g_\rho(r, \epsilon, e)$ $[r < \rho' < \rho]$, there follows that

(12) *The groups* $g_\rho(r, 1, e)$ $(\rho > r)$ *contain as a subgroup either the symmetric group* $\Pi_{\rho!}$, *or the alternating group* $\Pi_{\rho!/2}$, *of the variables* $x_1, \cdots, x_\rho$ *according as* $e = -1$ *or* $e = 1$.

For $e = -1$, this was already known in connection with the regular Cremona groups.

7. **Finite and infinite groups $g_\rho(\alpha)$ and $g_\rho(r, \epsilon, e)$.** There is a very simple geometrical criterion which separates the finite and infinite cases of the groups under consideration, namely:
(1) *If in the space $S_\rho(x_0, x_1, \cdots, x_\rho)$, the linear space $L = 0$ cuts the quadric $Q = 0$ in a quadric without real points, then the group g_ρ is finite; otherwise, it is infinite.*

To apply this criterion we write $Q = 0$, $L = 0$, using the notation of **1** (6) in the form $x_1^2 + \cdots + x_\rho^2 = \delta/\beta$, $(x_1 + \cdots + x_\rho)/\sqrt{\rho} = r\delta/(\alpha - 1)\sqrt{\rho}$. In this metric form the condition given in (1) is obviously $\delta/\beta < r^2\delta^2/\rho(\alpha - 1)^2$. This reduces by virtue of $r\delta\beta = \alpha^2 - 1$ and $\alpha > 1$ to the form

(2) $$\rho(\alpha - 1) < r(\alpha + 1).$$

We divide the cases as before into

(2a) $$r = 1, \quad \rho < (\alpha + 1)/(\alpha - 1);$$

(2b) $$r > 1, \quad e = -1, \quad \rho < r^2\epsilon/(r\epsilon - 2);$$

(2c) $$r > 1, \quad e = 1, \quad \rho < (r\epsilon + 2)/\epsilon.$$

In the case (2a) with $\rho \geqq 2$, $\alpha \geqq 2$ and $\rho < 1 + 2/(\alpha - 1)$, there is only one solution $\rho = 2$, $\alpha = 2$.

In the case (2b) with $\rho > r > 1$, $\epsilon \geqq 1$ and $\rho < r + 2r/(r\epsilon - 2)$, since $\rho - r \geqq 1$, then $1 < 2r/(r\epsilon - 2)$, or $r(\epsilon - 2) < 2$. Hence $\epsilon = 2$ or $\epsilon = 1$. If $\epsilon = 2$, $\rho < r + 1 + 1/(r - 1)$ or $\rho = r + 1$. If $\epsilon = 1$, the case $r = 2$ being excluded, $\rho < r + 2 + 4/(r - 2)$. Thus for $r = 3$, $\rho < 9$; for $r = 4$, $\rho < 8$; for $r = 5$, $\rho < 9$; and for $r > 5$, $\rho = r + 1, r + 2$.

In the case (2c) with $\rho > r > 1$, $\epsilon \geqq 1$, and $\rho < r + 2/\epsilon$, there is only one solution $\epsilon = 1$, $\rho = r + 1$.
(3) *The groups $g_\rho(\alpha)$, $g_\rho(r, \epsilon, e)$ are finite in the following cases*: (a) $g_\rho(\alpha) = g_2(2)$; (b) $g_\rho(r, \epsilon, -1)$ *for* $\epsilon = 1, 2$ *and* $\rho = r + 1$; *for* $\epsilon = 1$ *and* $\rho = r + 2$; *for* $\epsilon = 1$, $r = 3$, *and* $\rho = 6, 7, 8$; *for* $\epsilon = 1$, $r = 4$ *and* $\rho = 7$; *for* $\epsilon = 1$, $r = 5$ *and* $\rho = 8$; (c) $g_\rho(r, \epsilon, 1)$ *for* $\epsilon = 1$, $\rho = r + 1$. *In all other cases the groups are infinite.*

In this section we prove that the cases mentioned are infinite, leaving a discussion of the finite cases for the next section. We observe first that the cases (b) for $\epsilon = 1$ are known in connection with the regular Cremona transformations. It has been proved [cf. [1]II, §5, pp. 375, 377] that $g_9(3, 1, -1)$ and $g_8(4, 1, -1)$ are infinite. Hence $g_{9+k}(3, 1, -1)$ and $g_{8+l}(4, 1, -1)$ also are infinite [cf. **2** (6)]. According to **3** (3), their associated groups, $g_{9+k}(6 + k, 1, -1)$ and $g_{8+l}(4 + l, 1, -1)$ are infinite. Again by the use of **2** (6), we see that $g_{9+k+m}(6 + k, 1, -1)$ and $g_{8+l+n}(4 + l, 1, -1)$ are infinite. But these cover all the cases (b) for $\epsilon = 1$ except those asserted to be finite.

There remain to be considered only the cases

$$(d_1)\quad g_\rho(\alpha) = g_3(2); \qquad (d_2)\quad g_\rho(\alpha \geqq 3); \qquad (d_3)\quad g_{r+2}(r, 2, -1);$$

$$(d_4)\quad g_{r+1}(r, \epsilon > 2, -1); \qquad (d_5)\quad g_{r+2}(r, 1, 1); \qquad (d_6)\quad g_{r+1}(r, \epsilon > 1, 1).$$

In each of these cases we have selected the smallest value of ρ which yields an infinite group, the larger values being accounted for by **2** (6).

In the case (d_1), with $g_3(2)$ and its generators defined as in **2** (3), it is easy to verify that the form

$$M_k \equiv (3k-1)^2 x_0 - (3k-1)(3k-2)x_1 - 3k(3k-2)x_2 - 3k(3k-1)x_1$$

is transformed by $(I_1I_2I_3)^2$ into M_{k+1}, whence $(I_1I_2I_3)^2$ has not a finite period, and $g_3(2)$ is infinite.

In the case (d_2) with $a = (\alpha - 1) \geqq 2$, we define sequences of polynomials P_k, C_k [cf. [4]p. 111] by the recursion formula

$$(4)\qquad \begin{aligned} P_k &= a\,P_{k-1} - P_{k-2}, \qquad P_{-1} = 0, \qquad P_0 = 1, \qquad P_1 = a; \\ C_k &= P_k + P_{k-1}. \end{aligned}$$

As immediate consequences of these definitions we find that

$$(5)\qquad \begin{aligned} \alpha P_{k-1} - C_{k-1} &= P_k, \qquad (\alpha^2 - 1)P_{k-1} - \alpha C_{k-1} = C_{k+1}, \\ \alpha P_{k-1} - C_k &= P_{k-2}, \qquad (\alpha^2 - 1)P_{k-1} - \alpha C_k = C_{k-2}. \end{aligned}$$

If then we define the forms

$$(6)\qquad D_k = P_{k-1}x_0 - C_kx_1 - C_{k-1}x_2, \qquad E_k = P_{k-1}x_0 - C_{k-1}x_1 - C_kx_2,$$

it is easy to verify from (5) that I_1 in $g_2(\alpha)$ interchanges E_{k-1}, D_k, and I_2 interchanges D_{k-1}, E_k whence I_2I_1 transforms D_k into D_{k+2} with respective leading coefficients P_{k-1}, P_{k+1}. From (4) it follows that $P_k - P_{k-1} \geqq P_{k-1} - P_{k-2}$, whence, since $P_1 - P_0 = a - 1 \geqq 1$, the polynomials P_k continue to increase and I_2I_1 has not a finite period. Thus $g_2(\alpha \geqq 3)$ is infinite.

In the case (d_3), $g_{r+2}(r, 2, -1)$ is associated with, and isomorphic to, $g_{r+2}(2, 2, -1)$. Since $r \geqq 2$, $r + 2 \geqq 4$. In any case $g_{r+2}(2, 2, -1)$ contains subgroups $g_4(2, 2, -1)$ and it is sufficient to show that $g_4(2, 2, -1)$ is infinite. Consider the form

$$J_k \equiv \binom{2k+1}{2} x_0 - (2k^2 - 1)\,x_1 - 2k^2\,x_2 - 4\binom{k+1}{2}(x_3 + x_4).$$

It is easy to verify that the element $I_{12}I_{34}$ of $g_4(2, 2, -1)$ transforms J_k into J_{k+1}. Hence this element is aperiodic and $g_4(2, 2, -1)$ is infinite.

In the case (d_4), $g_{r+1}(r, \epsilon > 2, -1)$ is associated with $g_{r+1}(\alpha \geqq 2)$ [cf. §3 (4)]. Since $r \geqq 2$, $g_{r+1}(\alpha)$ is either $g_3(\alpha)$ or contains subgroups $g_3(\alpha)$. But, according to cases (d_1) or (d_2), $g_3(\alpha \geqq 2)$ is infinite.

[4] S. F. Barber, *Planar Cremona transformations*, Amer. Jour., vol. 56 (1934), pp. 109–121.

In the case (d_5) of $g_{r+2}(r, 1, 1)$ denote the generator $I_{i_1 \cdots i_r}$ by $J_{i_{r+1} i_{r+2}}$. Since $r \geqq 2$, $r + 2 \geqq 4$, and generators $J_{r-1, r}$, $J_{r+1, r+2}$ with non-overlapping indices exist. Again it is not difficult to verify that the form

$$L_k \equiv \binom{4k}{2} [x_0 - x_1 - \cdots - x_{r-2}] - 8 k^2 x_{r-1} - (8 k^2 - 1) x_r$$
$$- 4 \binom{2k}{2} (x_{r+1} + x_{r+2})$$

is transformed by $(J_{r-1, r} \cdot J_{r+1, r+2})^2$ into L_{k+1}. Hence this element of the group is aperiodic and the group is infinite.

In the case (d_6) of $g_{r+1}(r, \epsilon > 1, 1)$ let $I_{i_1 \cdots i_r} = J_{i_{r+1}}$. The linear form x_r transformed alternately by J_{r+1}, J_r yields the sequence

$$[x_0 - \epsilon(x_1 + \cdots + x_{r-1})] - (\epsilon + 1)x_r$$
$$(\epsilon + 1)[x_0 - \epsilon(x_1 + \cdots + x_{r-1})] - (\epsilon + 1)x_r - \epsilon(\epsilon + 2)x_{r+1},$$
$$\epsilon(\epsilon + 1)[x_0 - \epsilon(x_1 + \cdots + x_{r-1})] - (\epsilon + 1)(\epsilon^2 + \epsilon - 1)x_r - \epsilon(\epsilon + 2)x_{r+1},$$
$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots.$$

These forms have the type

$$\text{(7)} \qquad \begin{gathered} a[x_0 - \epsilon(x_1 + \cdots + x_{r-1})] - bx_r - cx_{r+1}, \\ (\epsilon + 2)a - b - c = 1. \end{gathered}$$

If (7) is transformed by J_{r+1}, it acquires coefficients a', b', c', where

$$\text{(8)} \quad J_{r+1}\colon a' = (1 + \epsilon)a - b, \qquad b' = \epsilon(\epsilon + 2)a - (1 + \epsilon)b, \qquad c' = c.$$

The interchange in this of b, c and of b', c' gives the effect of J_r upon the form (7). But this J_r and J_{r+1} in (8) are the generators I_2, I_1 of a $g_2(\alpha)$, where $\alpha = 1 + \epsilon \geqq 3$ and in which also $\beta = 1$, $\delta = \alpha^2 - 1$. Thus, according to the result in case (d_2), the effect of $J_r J_{r+1}$ upon forms (7) is aperiodic, whence this element $J_r J_{r+1}$ is aperiodic and $g_{r+1}(r, \epsilon > 1, 1)$ is infinite.

8. **Finite groups $g_\rho(\alpha)$ and $g_\rho(r, \epsilon, e)$.** In this section we verify the finiteness asserted in **7** (1), (3) and determine the nature of the finite groups. We take up the cases as they are listed in **7** (3).

With respect to the case (a), or $g_2(2)$, we have seen [cf. **6** (1) et seq.] that this is a dihedral group of order six.

With respect to the case (b) we observe that, apart from $g_{r+1}(r, 2, -1)$, all the instances have the form $g_\rho(r, 1, -1)$ and therefore have been found in connection with regular Cremona transformations. We list these in order.

The $g_{r+1}(r, 1, -1)$ $(r > 2)$ is a $g_{(r+2)!}$ isomorphic with the symmetric group on $r + 2$ things. This is the linear group associated with Moore's cross-ratio group of Cremona transformations [cf. [5]].

[5] E. H. Moore, *The cross-ratio group of n! Cremona transformations of order n − 3 in flat space of n − 3 dimensions*, Amer. Jour., vol. 22 (1900), pp. 279–291.

The $g_{r+2}(r, 1, -1)$ $(r > 2)$ is a $g_{(r+2)!\,2^{r+1}}$. It has an invariant Abelian $g_{2^{r+1}}$ whose factor group is isomorphic with the symmetric group. If $r = 2p$, it has an invariant g_2 whose factor group of order $(2p + 2)!\cdot 2^{2p}$ is isomorphic with the collineation group which transforms into itself the collineation group of order 2^{2p} induced on a hyperelliptic Kummer p-way by the addition of the 2^{2p} half periods. If $r = 2p - 1$, the group is a subgroup of the one just described [cf. [6]].

The self-associated [cf. [1]II, §5] $g_6(3, 1, -1)$ is isomorphic with the g_{51840} of the 27 lines on a cubic surface [cf. [1]II, §3].

The two associated groups $g_7(3, 1, -1)$, $g_7(4, 1, -1)$ each have an invariant g_2 whose factor group of order $8!\,36$ is isomorphic with the group of the double tangents of a planar quartic [cf. [1]II, §3].

The two associated groups $g_8(3, 1, -1)$, $g_8(5, 1, -1)$, each have an invariant g_2 whose factor group of order $10!\,96$ is isomorphic with the group of the tritangent planes of a space sextic of genus four on a quadric cone [cf. [1]II, §3].

In the last instance under case (b), according to **6** (7), the generating involutions, $J_1, \cdots, J_{r+1}$, are permutable and the finite $g_{r+1}(r, 2, -1)$ is abelian, of order 2^{r+1}, and type $(1, 1, \cdots, 1)$.

There remains only the case (c) of $g_{r+1}(r, 1, 1)$. The theorems **6** (8), (9), (10) indicate that this group is isomorphic with a symmetric $g_{(r+2)!}$. We make this proof more precise by using the set of $r + 2$ linear forms:

$$\begin{aligned} m_\infty &\equiv -(r+1)x_0 + (r+2)(x_1 + \cdots + x_{r+1}), \\ m_i &= x_0 - (r+2)x_i \qquad (i = 1, \cdots, r+1), \\ m_\infty &+ m_1 + \cdots + m_{r+1} \equiv 0. \end{aligned} \tag{1}$$

We find that J_{r+1} changes the sign of each of $m_1, \cdots, m_r$, and interchanges m_∞, m_{r+1} with a change of sign in each. Moreover, the transforms of $r + 1$ of these forms, and the invariance of L, defines the transformation. Hence $g_{r+1}(r, 1, 1)$ is isomorphic with the permutation $g_{(r+2)!}$ of these $r + 2$ forms. The form m_∞ is the form $-L'$ of **1** (7) formed for the symmetric element I used in **6** (9). The $g_{r+1}(r, 1, 1)$ does not contain the element I. If this were added as an additional generator, the doubled group with invariant $g_2 = 1I$ would contain the odd as well as the even permutations of $x_1, \cdots, x_{r+1}$.

The geometric criterion given in **7** (1) for finiteness recalls the theorem of Minkowski ([7]p. 185) that a positive quadratic form in n variables cannot admit more than $(2^{n+1} - 2)^n$ integral linear transformations. Hence it can admit only a finite group of such transformations. Yet this theorem can not be applied here directly. For if, say, x_0 is eliminated by using the invariant linear relation $L = 0$, our group is converted into one which has rational coefficients [cf. [1]II, §6].

[6] A. B. Coble, *A generalization of the Weddle surface*, etc., Amer. Jour., vol. 52 (1930), pp. 439–500.

[7] H. Minkowski, *Geometrie der Zahlen*, Leipzig, 1896.

9. **The generic element T of the group $g_\rho(\alpha)$.** We have in **4** taken this element of the group $g_\rho(\alpha)$ generated by I_i $(i = 1, \cdots, \rho)$,

$$(1) \qquad \begin{gathered} x_0' = \alpha x_0 - \beta x_i, \quad x_i' = \delta x_0 - \alpha x_i, \quad x_j' = x_j \quad (j = 1, \cdots, \rho; j \neq i). \\ \alpha \geqq 2, \quad \delta\beta = \alpha^2 - 1, \end{gathered}$$

in the form

$$(2) \qquad T: x_m' = \alpha_{m0}x_0 - \alpha_{m1}x_1 - \cdots - \alpha_{m\rho}x_\rho \quad (m = 0, 1, \cdots, \rho).$$

We have also obtained certain numerical properties [**4** (2)] of these coefficients, and certain quadratic and linear relations [**4** (3), (5)] satisfied by them. We seek to prove that, with one type of exception noted below, the coefficients α_{mn} $(m, n = 0, 1, \cdots, \rho)$ are positive integers or zero. If T could be attached to a type of Cremona transformation, this could be inferred from the fact that geometric multiplicities are positive or zero. We proceed then to define a type of Cremona transformation determined by T.

The ternary de Jonquières transformation of order n has a set of $2(n - 1)$ simple F-points and a single F-point of order $n - 1$. Its effect upon curves with order t_0, and multiplicities $y_1, \cdots, y_{2(n-1)}$ at the simple F-points, multiplicity t_1 at the $(n - 1)$-fold F-point, and multiplicities $t_2, \cdots, t_\rho$ at $\rho - 1$ further ordinary points is expressed by the equations:

$$(3) \qquad J_1: \begin{aligned} t_0' &= nt_0 - (y_1 + \cdots + y_{2(n-1)}) - (n-1)t_1, \\ y_i' &= t_0 - y_i - t_1 && (i = 1, 2, \cdots, n-1), \\ t_1' &= (n-1)t_0 - (y_1 + \cdots + y_{2(n-1)}) - (n-2)t_1, \\ t_j' &= t_j && (j = 2, \cdots, \rho). \end{aligned}$$

The t', y_i' are multiplicities at the inverse F-points of the transformation. If now we form products of such transformations, always taking the simple F-points of the last factor at the $2(n - 1)$ inverse F-points of the preceding product which arose earlier from the simple points of the factors of this product, and not at any time allowing more than ρ additional F-points to appear from the $(n - 1)$-fold F-points of the factors (i.e., forming products as though the $2(n - 1)$ F-points could be fixed for all factors), we then secure an aggregate of types of products whose effect upon curves is given by the aggregate of elements of the group generated by $J_1, \cdots, J_\rho$, each J being formed like J_1 in (3).

Since $y_1, \cdots, y_{2(n-1)}$ occur symmetrically in all of the J's, let

$$(4) \qquad \sigma = y_1 + y_2 + \cdots + y_{2(n-1)}.$$

Then (3) takes the form

$$(5) \qquad \begin{aligned} t_0' &= nt_0 - \sigma - (n-1)t_1, \\ \sigma' &= 2(n-1)t_0 - \sigma - 2(n-1)t_1, \\ t_1' &= (n-1)t_0 - \sigma - (n-2)t_1, \\ t_j' &= t_j && (j = 2, \cdots, \rho). \end{aligned}$$

This also is an involutorial transformation on the variables σ, t. We now make the change of variable

$$2x = -\sigma + 2t_0, \qquad x + x_j = t_j \qquad (j = 0, 1, \cdots, \rho). \tag{6}$$

Then all of the generators J take the typical form

$$\begin{aligned} x_0' &= (n-2)x_0 - (n-1)x_1, \\ x_1' &= (n-3)x_0 - (n-2)x_1, \\ x' &= x, \quad x_2' = x_2, \quad \cdots, \quad x_\rho' = x_\rho. \end{aligned} \tag{7}$$

But (7) is the generator I_1 of $g_\rho(\alpha)$ amplified by $x' = x$, where

$$\alpha = n - 2, \quad \delta = n - 3, \quad \beta = n - 1, \quad \text{and} \quad \delta\beta = \alpha^2 - 1. \tag{8}$$

Hence

(9) *The generic type of ternary Cremona transformation which is a product, formed as indicated above, of de Jonquières transformations of order n is isomorphic with the generic element* (2) *of the group* $g_\rho(n-2)$.

Suppose that the generic element T of $g_\rho(n-2)$ with constants (8) is given as in (2) with the necessary conditions on its coefficients as determined in **4** (2), (7), (8), namely

$$\begin{aligned} &\alpha_{00}^2 - 1 = a_0(n-3)(n-1), \qquad \alpha_{i0} = (n-3)d_i, \\ &\qquad\qquad \alpha_{0j} = (n-1)b_i \qquad (i, j = 1, \cdots, \rho), \\ &\alpha_{00} - \Delta = \mu(n-1), \qquad \mu - \Sigma_j b_j = 2\lambda. \end{aligned} \tag{10}$$

We seek to determine the corresponding type of ternary Cremona transformation. From (6) and (2) we find that

$$\begin{aligned} t_0' &= x' + x_0 = x + \alpha_{00}x_0 - \Sigma_j\alpha_{0j}x_j = x + \alpha_{00}(t_0 - x) - \Sigma_j\alpha_{0j}(t_j - x) \\ &= x[1 - \alpha_{00} + \Sigma_j\alpha_{0j}] + \alpha_{00}t_0 - \Sigma_j\alpha_{0j}t_j \\ &= \tfrac{1}{2}(-\sigma + 2t_0)[1 - \alpha_{00} + \Sigma_j\alpha_{0j}] + \alpha_{00}t_0 - \Sigma_j\alpha_{0j}t_j. \end{aligned}$$

On substituting for α_{00} and α_{0j} from (10), we have for t_0', and similarly for the other variables,

$$\begin{aligned} t_0' &= \{1 + (n-1)\Sigma_j b_j\}t_0 - \{\Sigma_j b_j\}(y_1 + \cdots + y_{2(n-1)}) - (n-1)\Sigma_j b_j t_j, \\ y_i' &= \{\Sigma_j b_j\}t_0 - \tfrac{1}{2}\{\Sigma_j b_j - \mu\}(y_1 + \cdots + y_{2(n-1)}) + \Delta y_i - \Sigma_j b_j t_j \\ &\qquad\qquad [i = 1, \cdots, 2(n-1)], \\ t_k' &= \{(n-1)d_k\}t_0 - d_k(y_1 + \cdots + y_{2(n-1)}) - \Sigma_{j'}\alpha_{kj'}t_j \\ &\qquad\qquad [k = 1, \cdots, \rho]. \end{aligned} \tag{11}$$

This value of y_i' is determined indirectly from σ'. For, $-\sigma' + 2t = -\sigma + 2t_0$, whence

$$\sigma' = \{2(n-1)\Sigma_j b_j\}t_0 - (2\Sigma_j b_j - 1)\sigma - 2(n-1)\Sigma_j b_j t_j.$$

Now σ' is obtained from $y'_1 + \cdots + y'_{2(n-1)}$, and the values of the y''s are the same except in the matrix of coefficients of $y_1, \cdots, y_{2(n-1)}$. This square matrix has like elements $-m$ except in the principal diagonal where they are $-(m \pm 1)$. On comparing the σ' arising from such rows with that just given, and noting that

$$2\Sigma_j b_j - 1 = 2(n-1)\{\tfrac{1}{2}(\Sigma_j b_j - \mu)\} - \Delta,$$

we obtain the y'_i given in (11). Since $\Sigma_j b_j - \mu$ is even [cf. **4** (8)], the coefficients in (11) are integral.

(12) *The transformation* (11) *defined by the coefficients α_{ij} of T in $g_\rho(n-2)$ represents a geometrically existent ternary Cremona transformation, the product of a set of de Jonquières transformations. It is formed from generators J in* (3) *precisely as T is formed from generators I.*

Since (11) is attached to a geometrically existent transformation, its coefficients have a variety of properties which can be transferred at once to the coefficients of T which are found in (11). We recapitulate some of these:

(13) (a) *The coefficients α_{ij} of the generic transformation T of $g_\rho(\alpha = n-2)$ with $\delta = n-3$, $\beta = n-1$ are positive integers or zero with the exception that if $\alpha_{ij} = -1$ $(i > 0, j > 0)$, then every coefficient in the same row and column as a_{ij} is zero.* (b) *With every set of κ equal numbers b_j there is associated a set of κ equal numbers d_i. The κ columns and κ rows of the matrix of coefficients thus isolated have in common a square matrix M_κ whose elements are all $-m$ except that in each row and column of M_κ there is one element $-(m \pm 1)$. Otherwise the κ columns are identical and the κ rows are identical.* (c) *For a particular element α_{ij} $(i > 0, j > 0)$, the inequality $(n-1)(b_j + d_i) \leqq 1 + (n-1)\Sigma_j b_j + \alpha_{ij}$ is valid.*

This last inequality arises from the known inequality (cf. [1]II, p. 368), $\sigma_j + \rho_i \leqq m + \alpha_{ij}$, which occurs in connection with a ternary transformation of order m which possesses an F-point of order ρ_i and of multiplicity α_{ij} on a P-curve of order σ_j.

We have just used the ternary transformations to obtain properties of the group $g_\rho(\alpha)$, but clearly the reaction is mutual, as is expressed in (12). Thus one may read off from the rows and columns of coefficients in (11), the characteristics of homaloidal nets and of P-curves in terms of the rows and columns of the comparatively simple elements of the group $g_\rho(\alpha)$. We may then expect to find applications of the groups $g_\rho(\alpha)$ and $g_\rho(r, \epsilon, e)$, not merely in connection with new space transformations and groups where they first were observed, but also in connection with aggregates of transformations already studied.

The underlying restriction $\alpha \geqq 2$, applied in the above account, yields $n - 2 \geqq 2$ or $n \geqq 4$. It is interesting to notice that in the excluded cases $n = 2, 3$ the products of de Jonquières transformations as used above may be regarded as the elements of a ternary Cremona group. For if the $2(n-1)$ simple F-points of the direct and inverse transformations could be superposed at positions $p_1, \cdots, p_{2(n-1)}$, fixed for all the generators, these generators would yield an actual group. Thus, if, for $n = 2$, p_1, p_2 are fixed at the circular points and

p_3, p_3' is variable, the group of inversions is determined. If, for $n = 3$, the four simple points of the cubic transformation, both direct and inverse, are fixed at $p_1, \cdots, p_4$, then the double F-points must coincide at O_1, since the five direct, and five inverse, F-points are projective. The generator J_{O_1} is then the projection of each conic of the pencil on $p_1, \cdots, p_4$ into itself from O_1. A quadratic involution with F-triangle at p_2, p_3, p_4 and fixed point p_1 converts the conics into lines on p_1, and the lines on O_1 into conics on p_2, p_3, p_4, P_1. Thus our generators are transformed into cubic involutions with a fixed double F-point, three fixed simple F-points and one variable F-point. They generate an infinite de Jonquières group with a pencil of invariant lines through p_1.

For $n \geqq 4$ the two sets of direct and inverse simple points of the de Jonquières transformation cannot lie in superposed position.

10. **Inequalities satisfied by the integral coefficients of the elements of $g_\rho(\alpha)$.** In this section we take $g_\rho(\alpha)$ with $\delta = 1$, $\beta = \alpha^2 - 1$. We shall be concerned primarily with those linear forms which are respectively the conjugates of x_0, and the conjugates of x_i, under $g_\rho(\alpha)$, i.e., with the rows of the various elements in $g_\rho(\alpha)$. It is convenient to use also the contragredient form $g_\rho(\alpha)_\xi$ of $g_\rho(\alpha)_x$ which is defined by the following contragredient invariant:

$$\xi_0' x_0' - (\xi_1' x_1' + \cdots + \xi_\rho' x_\rho') = \xi_0 x_0 - (\xi_1 x_1 + \cdots + \xi_\rho x_\rho). \tag{1}$$

Then the contragredient generators of $g_\rho(\alpha)_x$ and $g_\rho(\alpha)_\xi$ are

$$I_1(x): \begin{array}{l} x_0' = \alpha x_0 - (\alpha^2 - 1)x_1, \\ x_1' = x_0 - \alpha x_1; \end{array} \qquad I_1(\xi): \begin{array}{l} \xi_0' = \alpha\xi_0 - \xi_1, \\ \xi_1' = (\alpha^2 - 1)\xi_0 - \alpha\xi_1. \end{array} \tag{2}$$

The invariants of the two groups are respectively:

$$\begin{aligned} &L(x) = x_0 - (\alpha - 1)(x_1 + \cdots + x_\rho), \\ &\qquad\qquad L(\xi) = (\alpha + 1)\xi_0 - (\xi_1 + \cdots + \xi_\rho), \\ &Q(x) = x_0^2 - (\alpha^2 - 1)(x_1^2 + \cdots + x_\rho^2), \\ &\qquad\qquad Q(\xi) = (\alpha^2 - 1)\xi_0^2 - (\xi_1^2 + \cdots + \xi_\rho^2). \end{aligned} \tag{3}$$

A set of values c of the x's will be called a *characteristic* $C(x)$ $(x = c)$, and this set will often be indicated by a form $c_0\xi_0 - (c_1\xi_1 + \cdots + c_\rho\xi_\rho)$; a set of values γ of the ξ's will be called a *characteristic* $C(\xi)$ $(\xi = \gamma)$, and again this set will often be indicated by a form $\gamma_0 x_0 - (\gamma_1 x_1 + \cdots + \gamma_\rho x_\rho)$.

The characteristics $C(x)$ indicated by ξ_0 and its conjugates will be called *characteristics* $H(x)$. A list of the early conjugates follows:

$$\begin{aligned} H &= \xi_0, \\ H_i &= \{1 + (\alpha - 1)\}\xi_0 - \xi_i, \\ H_{ji} &= \{1 + (\alpha - 1)(\alpha + 1)\}\xi_0 - \alpha\xi_i - \xi_j, \end{aligned}$$

$$
\begin{aligned}
H_{iji} &= \{1 + (\alpha - 1)\alpha^2\}\xi_0 - (\alpha^2 - \alpha)\xi_i - \alpha\xi_j, \\
H_{kji} &= \{1 + (\alpha - 1)(\alpha^2 + \alpha + 1)\}\xi_0 - \alpha^2\xi_i - \alpha\xi_j - \xi_k, \\
H_{jiji} &= \{1 + (\alpha - 1)(\alpha^3 - \alpha^2 - \alpha + 1)\}\xi_0 - (\alpha^3 - 2\alpha^2 + 1)\xi_i, \\
&\qquad\qquad - (\alpha^2 - \alpha)\xi_j. \\
H_{jkji} &= \{1 + (\alpha - 1)(\alpha^3 + 1)\}\xi_0 - (\alpha^3 - \alpha^2 + 1)\xi_i - (\alpha^2 - \alpha)\xi_j - \alpha\xi_k. \\
H_{kiji} &= \{1 + (\alpha - 1)(\alpha^3 + 1)\}\xi_0 - (\alpha^3 - \alpha^2)\xi_i - \alpha^2\xi_j - \xi_k, \\
H_{ikji} &= \{1 + (\alpha - 1)(\alpha^3 + \alpha^2)\}\xi_0 - (\alpha^3 - \alpha)\xi_i - \alpha^2\xi_j - \alpha\xi_k, \\
H_{lkji} &= \{1 + (\alpha - 1)(\alpha^3 + \alpha^2 + \alpha + 1)\}\xi_0 - \alpha^3\xi_i - \alpha^2\xi_j - \alpha\xi_k - \xi_l, \\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots .
\end{aligned} \tag{4}
$$

It is clear that

(5) *The set of conjugates* (4) *of* ξ_0 *under* $g_\rho(\alpha)_\xi$ *have for polars as to* $Q(x)$ *the set of conjugates of* x_0 *under* $g_\rho(\alpha)_x$. *The form* $H_{ji\dots}$ *is the transform of* H *by the product* $I_j(\xi)\cdot I_i(\xi)\cdots$. *The coefficients* $h_0, h_1, \cdots, h_\rho$ *of a particular form* $h_0\xi_0 - (h_1\xi_1 + \cdots + h_\rho\xi_\rho)$ *satisfy the relations* $Q(h) = 1, L(h) = 1$.

The characteristics $C(\xi)$ indicated by x_j and its conjugates will be called *characteristics* $P(\xi)$. A list of the early conjugates follows:

$$
\begin{aligned}
p_j &= x_j, \\
P_j &= x_0 - \alpha x_j, \\
P_{jk} &= \alpha x_0 - (\alpha^2 - 1)x_k - \alpha x_j, \\
P_{jkj} &= (\alpha^2 - \alpha)x_0 - \alpha(\alpha^2 - \alpha - 1)x_j - (\alpha^2 - 1)x_k, \\
P_{jkl} &= \alpha^2 x_0 - \alpha(\alpha^2 - 1)x_l - (\alpha^2 - 1)x_k - \alpha x_j, \\
P_{jkjk} &= (\alpha^3 - 2\alpha^2 + 1)x_0 - (\alpha^4 - 2\alpha^3 - \alpha^2 + 2\alpha)x_k - \alpha(\alpha^2 - \alpha - 1)x_j, \\
P_{jkjl} &= \alpha^2(\alpha - 1)x_0 - (\alpha^2 - \alpha)(\alpha^2 - 1)x_l - \alpha(\alpha^2 - \alpha - 1)x_j - (\alpha^2 - 1)x_k, \\
P_{jklk} &= (\alpha^3 - \alpha^2 + 1)x_0 - (\alpha^4 - \alpha^3 - \alpha^2 + \alpha)x_k - \alpha(\alpha^2 - 1)x_l \\
&\qquad\qquad - (\alpha^2 - 1)x_k, \\
P_{jklj} &= (\alpha^3 - \alpha)x_0 - (\alpha^4 - 2\alpha^2)x_j - \alpha(\alpha^2 - 1)x_l - (\alpha^2 - 1)x_k, \\
P_{jklm} &= \alpha^3 x_0 - \alpha^2(\alpha^2 - 1)x_m - \alpha(\alpha^2 - 1)x_l - (\alpha^2 - 1)x_k - \alpha x_j, \\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots .
\end{aligned} \tag{6}
$$

Again it is clear that

(7) *The set of conjugates* (6) *of* x_j *under* $g_\rho(\alpha)_x$ *are such that* $P_{jk\dots}$ *is the transform of* x_j *by the product* $I_j(x)\cdot I_k(x)\cdots$. *The coefficients* $\pi_0, \pi_1, \cdots, \pi_\rho$ *of a particular one of these forms,* $\pi_0x_0 - (\pi_1x_1 + \cdots + \pi_\rho x_\rho)$, *satisfy the relations* $L(\pi) = 1$, $Q(\pi) = -1$.

We prove first the theorem:

(8) *Any characteristic* $(h_0; h_1 \cdots h_\rho)$ *for which* $L(h) = 1$, $Q(h) = 1$, *and for which* $h_j \geqq 0$ $(j = 1, \cdots, \rho)$, *when arranged so that* $h_1 \geqq h_2 \geqq h_3 \geqq \cdots \geqq h_\rho$, *satisfies the inequalities*

$$0 < h_0 < (\alpha + 1)h_1$$

except in the one instance $(1; 0 \cdots 0)$ *given by* H *in* (4).

For, from $L(h) = 1$, $\alpha \geqq 2$ and $h_j \geqq 0$, it follows that $h_0 > 0$. If $h_0 = 1$, it follows from $h_1 + \cdots + h_\rho = 0$ that $h_j = 0$ $(j = 1, \cdots, \rho)$. This yields the excluded characteristic $(1; 0 \cdots 0)$. If $h_0 > 1$, $h_1 > 0$. If then $h_2 = \cdots = h_\rho = 0$, $L(h) = Q(h) = 1$ become $(\alpha - 1)h_1 = h_0 - 1$, $(\alpha^2 - 1)h_1^2 = h_0^2 - 1$. By division, $(\alpha + 1)h_1 = h_0 + 1$, whence $(h) = (\alpha; 10 \cdots 0)$, which satisfies (8). If $h_2 > 0$, we use the inequality $h_1(h_2 + \cdots + h_\rho) \geqq h_1^2 + h_2^2 + \cdots + h_\rho^2$. This, combined with $L(h) = Q(h) = 1$, yields $[(h_0 - 1) - (\alpha - 1)](\alpha + 1)h_1 \geqq (h_0^2 - 1) - (\alpha^2 - 1)h_1^2$. Hence $(h_0 - 1)(\alpha + 1)h_1 \geqq h_0^2 - 1$. Since $h_0 > 1$, $(\alpha + 1)h_1 \geqq h_0 + 1$, or $h_0 < (\alpha + 1)h_1$.

We prove further that

(9) *Every characteristic* $(h; h_1 \cdots h_\rho)$ *which satisfies* $L(h) = Q(h) = 1$, *and which satisfies the system* (*infinite except when* $\rho = 2$, $\alpha = 2$) *of inequalities* $P_{jk} \cdots (x = h) \geqq 0$ *obtained from the forms* (6), *can be reduced by transformations in* $g_\rho(\alpha)_z$ *to the characteristic* $(1; 0 \cdots 0)$.

For, since (h) satisfies $x_j \geqq 0$, the $h_1, \cdots, h_\rho$ are positive or zero, whence h_0 is positive. According to (8) there is an h_j such that $h_0 < (\alpha + 1)h_j$. Then $I_j(x)$ applied to (h) yields a characteristic (h') for which $h_0' < h_0$ and which also satisfies the inequalities $h_j' \geqq 0$, since these inequalities for h' arise from inequalities in the set (4) satisfied by (h). The element h_0' can therefore be continually reduced until it reaches 1, in which case the remaining h's are all zero, since they still satisfy $x_j \geqq 0$.

It is probable that all of these inequalities are independent. For example, when $\alpha = 2$, there is a case $(h) = (11; 61^4)$ which satisfies $x_j \geqq 0$ but which does not satisfy $P_j \geqq 0$. There is also a case $(h) = (20; 971^3)$ which satisfies $x_j \geqq 0$ and $P_j \geqq 0$, but which does not satisfy $P_{jk} \geqq 0$.

A list of characteristics for which $L = Q = 1$ and $x_j \geqq 0$, for this case $\alpha = 2$ and complete up to $n = 17$ is as follows:

$(1; 0)$, $(2; 1)$, $(4; 2, 1)$, $(5; 2^2)$, $(8; 4, 2, 1)$, $(10; 4^2, 1)$, $(10; 5, 2^2)$, $(11; 6, 1^4)$*,

$(13; 6, 4, 2)$, $(13; 7, 2, 1^3)$*, $(14; 6, 5, 2)$, $(16; 8421)$, $(17; 9321^2)$*,

those marked * not being reducible to $(1; 0 \cdots 0)$.

(10) *If a characteristic* $(h) = (h_0; h_1 \cdots h_\rho)$ *is reducible in* $g_\rho(\alpha)_z$, *when* $\alpha \geqq 3$, *to* $(1; 0 \cdots 0)$, *it satisfies the following system of inequalities*:

(I) $\alpha h_1 \leqq h_0 < (\alpha + 1)h_1$; (II) $h_1 > h_2 > \cdots > h_k > h_{k+1} = 0 = h_{k+2} = \cdots = h_\rho$;

(III) $h_0 > (\alpha + 1)h_2$; (IV) $h_0 - \alpha h_2 > h_1$.

For if (h) is reducible to $(1; 0 \cdots 0)$ in $g_\rho(\alpha)_z$, then conversely $(1; 0 \cdots 0)$ can be transformed into (h) by a sequence of involutions I_x. Such a set of transforms, $h_0\xi_0 - (h_1\xi_1 + \cdots + h_\rho\xi_\rho)$, of $(1; 0 \cdots 0)$ by sequences of four or fewer involutions is found in (4). The statements of (10) are satisfied by all of these transforms, and furthermore the largest h_i $(i = 1, \cdots, \rho)$ is that which arises from the involution last used. Suppose then that these statements remain true for transforms under all sequences of n or fewer involutions. If (h_n) is such a transform, we apply an additional involution to obtain a transform

$$(h'_{n+1}) = (\alpha h_0 - (\alpha^2 - 1)h_j;\ h_0 - \alpha h_j,\ h_1, \cdots, h_{j-1}, h_{j+1}, \cdots, h_\rho),$$

where $j \geqq 2$ else (h'_{n+1}) is an (h_{n-1}). Then the inequalities (II) are satisfied by (h'_{n+1}), since (II) and (IV) are satisfied by (h_n). The second part of the inequality (I) is then a consequence of (8). The first part of the inequality (I), $h'_0 \geqq \alpha h'_1$, reduces to $h_j \geqq 0$. The inequality (IV) is an immediate consequence of $L(h') = 1$ which can be written in the form

$$h'_0 - \alpha h'_2 = h'_1 + 1 + [(\alpha - 3)h'_1 + (\alpha - 1)(h'_3 + \cdots + h'_\rho) + (h'_1 - h'_2)].$$

Finally, the inequality (III) is a consequence of (IV) and $h'_2 < h'_1$.

(11) *The generic conjugate of x_0 in $g_\rho(\alpha)$ $(\alpha \geqq 3)$ has the form $x'_0 = h_0x_0 - (\alpha^2 - 1)(h_1x_1 + \cdots + h_\rho x_\rho)$, where the numbers $h_0, \cdots, h_\rho$, when properly ordered, are those described in* (10).

This brings out the rather unlooked for result that when $\alpha \geqq 3$, none of the $h_1, \cdots, h_\rho$ are equal unless they are zero. Thus the matrices M_ϵ of **9** (13)(b) do not exist. However this is not true when $\alpha = 2$, as the list preceding (10) shows.

The conditions, finite in number, given in (10) do not replace the infinite number mentioned in (9). They include the first of these, such as $x_j \geqq 0$ and $P_j \geqq 0$, but they do not lead to $P_{jk} \geqq 0$ [cf. (6)].

We now consider the characteristics $(\pi_0; \pi_1 \cdots \pi_\rho)$ mentioned in (7), examples of which are given in (6). The first of these, $(0; -1, 0, \cdots, 0)$, is somewhat exceptional in that $\pi_0 = 0$ and a π_i is negative. This is the only solution of the equations $L(\pi) = 1$, $Q(\pi) = -1$ for which $\pi_0 = 0$. It is also clear that the second example, $(1; \alpha\, 0^{\rho-1})$, is the only solution of these equations for which $\pi_2, \cdots, \pi_\rho$ are zero. We now prove that

(12) *Except for the two particular cases just mentioned, every characteristic π which satisfies the equations $L(\pi) = 1$, $Q(\pi) = -1$, and for which $\pi_i \geqq 0$ $(i = 1, \cdots, \rho)$, when arranged so that $\pi_1 \geqq \pi_2 \geqq \cdots \geqq \pi_\rho$, satisfies the inequalities*:

$$(\alpha - 1)\pi_0 < \pi_1 < \alpha\pi_0; \qquad \pi_2 < (\alpha - 1)\pi_0 \text{ if } \alpha \geqq 3.$$

For from the inequality $\pi_1(\pi_2 + \cdots + \pi_\rho) \geqq \pi_2^2 + \cdots + \pi_\rho^2$, it follows immediately from $L(\pi) = +1$ and $Q(\pi) = -1$ that $\pi_0\pi_1(\alpha + 1) \geqq (\alpha^2 - 1)\pi_0^2 + \pi_1 + 1$, or $\pi_0\pi_1(\alpha + 1) > (\alpha^2 - 1)\pi_0^2$, or $\pi_1 > (\alpha - 1)\pi_0$. Let then $\pi_1 = (\alpha - 1)\pi_0 + k$ $(k > 0)$. Again from $Q(\pi) = -1$ we find that $2(\alpha - 1)\pi_0(\pi_0 - k) = k^2 - 1 + \pi_2^2 + \cdots + \pi_\rho^2$. Since $k > 0$ and $\pi_2 > 0$, then $k < \pi_0$, and $\pi_1 < \alpha\pi_0$. If now also $\pi_2 = (\alpha - 1)\pi_0 + k_1$, then

$$-(\alpha - 1)(\alpha - 3)\pi_0^2 - 2(\alpha - 1)\pi_0 k - 2(\alpha - 1)\pi_0 k_1 = k^2 - 1 + k_1^2 + \pi_3^2 + \cdots + \pi_\rho^2.$$

The right member is positive or zero. On the left the first term is zero or negative if $\alpha \geqq 3$, the second is negative, whence $k_1 < 0$, or $\pi_2 < (\alpha - 1)\pi_0$.

(13) *With the exception mentioned in* (12) *every characteristic* π, *for which* $(\pi_0 x_0 - \pi_1 x_1 + \cdots + \pi_\rho x_\rho)$ *is a conjugate of an* x_j *under* $g_\rho(\alpha)$, *satisfies, when* $\alpha \geqq 3$, *not only the equalities of* (12) *but also the additional inequalities*:

$$(\alpha^2 - 1)\pi_0 > \pi_1 + \alpha\pi_2; \qquad \pi_1 > \pi_2 > \cdots > \pi_k > \pi_{k+1} = 0 = \pi_{k+2} = \cdots = \pi_\rho.$$

It is clear that this is true of the members in the list (6) and that further the largest integer π_1 of the characteristic arises from the involution last applied. Let us then assume that (13) is true for all cases (π_n) which arise from x_j by applying n or fewer involutions. If then we apply an additional involution (which is not I_1, since the transform would then be a (π_{n-1}) for which (13) is true), the transformed characteristic is

$$(\pi') = (\alpha\pi_0 - \pi_j; (\alpha^2 - 1)\pi_0 - \alpha\pi_j, \pi_1, \cdots, \pi_{j-1}, \pi_{j+1}, \cdots, \pi_\rho) \qquad (j \geqq 2).$$

If $(\alpha^2 - 1)\pi_0 - \alpha\pi_j > \pi_1$, i.e., if the first inequality is satisfied by (π), then the remaining inequalities are satisfied by π'. But this first inequality is a consequence of $L(\pi) = 1$, $\alpha \geqq 3$, and $\pi_j \geqq 0$. For $(\alpha = 1)L(\pi) = (\alpha - 1)$ can be written as

$$(\alpha^2 - 1)\pi_0 - \alpha\pi_2 - \pi_1 = (\alpha + 1) + [(\alpha - 3)\pi_1 + (\pi_1 - \pi_2) + (\alpha - 1)(\pi_3 + \cdots + \pi_\rho)].$$

On the right $\alpha - 1 > 0$, and every other term is zero or positive, whence the first inequality is satisfied by π and also by (π').

Some of the inequalities of (10) and (13) fail in the case of $\alpha = 2$ because of the presence of periodic elements other than the generators.

If $\alpha \neq 2$, we have the theorem

(14) *The groups* $g_\rho(\alpha)$, $\alpha \geqq 3$ *contain no periodic elements other than the involutorial generators.*

For if $T = I_{k_1} I_{k_2} \cdots I_{k_n}$ is an element of $g_\rho(\alpha)$ naturally so written that no two successive subscripts k are alike, then it may happen that $k_n = k_1$. Then T is the transform of $T' = I_{k_2} \cdots I_{k_{n-1}}$ by I_k and T' has the same period as T. The determination of the period may again be simplified if $k_2 = k_{n-1}$. Proceeding in this way, we must eventually find that T is the transform of a generator of period two or that T is the transform of an element $T'' = I_{k_{1+s}} \cdots I_{k_{n-s}}$, where $1 + s \neq n - s$. Then $(T'')^j$ is a product of $j(n - 2s)$ generators such that no two successive factors are alike. But it is clear from (10) (III) that, as these successive factors are adjoined, the values of h_0 in the characteristic $x_0' = h_0 x_0 - (\alpha^2 - 1)(h_1 x_1 + \cdots + h_\rho x_\rho)$ steadily increase so that $(T'')^j$ can never be the identity.

UNIVERSITY OF ILLINOIS.

COMMUTATOR ALGEBRA OF A FINITE GROUP OF COLLINEATIONS

By Hermann Weyl

1. **Introduction.** In Chapter V, §§2–4 of my book *The Theory of Groups and Quantum Mechanics* (English edition, London, 1931) I gave an elementary account of the decomposition of tensor space into subspaces invariant under the algebra of "symmetric" transformations. The treatment was based upon the reciprocity between the ring of a finite group γ whose elements s induce certain linear operators $\mathbf{s}$ in a given vector space $\mathfrak{R}$ and the algebra $\mathfrak{A}$ of those transformations A in $\mathfrak{R}$ that commute with all operators $\mathbf{s}$. It was immaterial that $\mathfrak{R}$ was the manifold of all tensors of a certain rank f in an underlying vector space and γ the symmetric group of all $f!$ permutations operating on the f indices or arguments of the tensors. In many respects the group ring stands in a simpler relationship to this commutator algebra $\mathfrak{A}$ than to the enveloping algebra $\mathfrak{B}$ of the operators $\mathbf{s}$, and therefore it seems desirable to discuss both sides in a direct way rather than to rely upon the general theory of a matric algebra and its commutator algebra (cf. in this regard the observations in the concluding section).

A *number field* k may be given in which all numbers which occur are supposed to lie. We consider vectors

$$f = (f_1, \cdots, f_n)$$

in an n-space $\mathfrak{R}$ whose components f_i are numbers in k. When the result of a linear transformation $U = \|u_{ik}\|$ on f is denoted by Uf, the components of f are to be written in a *column*. A *collineation* $f \to \mathfrak{f}$ in the projective $(n - 1)$-space based upon the number field k is a linear transformation U of the coördinates f_i combined with an automorphism s: $\alpha \to \alpha^s$ of k:

$$\mathfrak{f}_i = \sum_{k=1}^{n} u_{ik} f_k^s \qquad \text{or} \qquad \mathfrak{f} = Uf^s.$$

Representations of a finite group by operators of this generalized type involving automorphisms of the reference field were studied with remarkable success in a recent paper by Messrs. I. Nakayama and K. Shoda (Jap. Jour. Math., vol. 12 (1936), pp. 109–122). The same step will here be carried out with respect to the commutator algebra.

2. **The group ring.** The situation we are concerned with may be described thus. Given a *finite group* γ of order h; if k is of prime characteristic, that prime shall not be a divisor of h. To each element s of γ corresponds an automorphism $\alpha \to \alpha^s$ of k such that

$$(\alpha^s)^t = \alpha^{ts}.$$

Received March 1, 1937.

A *collinear representation* of degree n of γ associates with each s an operator $f \to \bar{f}$ in the n-space $\mathfrak{R}$ of the form

$$\bar{f}_i = \sum_k u_{ik}(s) f_k^s \qquad (i, k = 1, \cdots, n) \tag{2.1}$$

in such a way that composition of collineations reflects the composition of group elements. This is expressed by the following equation for the matrix $U(s) = \| u_{ik}(s) \|$:

$$U(st) = U(s)U^s(t) \qquad (s, t \text{ in } \gamma). \tag{2.2}$$

In particular,

$$U(s)U^s(s^{-1}) = E \qquad \text{or} \qquad U^s(s^{-1}) = U^{-1}(s). \tag{2.3}$$

(2.1) shall be indicated briefly by

$$\bar{f} = \mathbf{s}f. \tag{2.4}$$

We introduce the "quantities" of the *group ring* ρ by the formal sum

$$\mathbf{a} = \sum_s a(s) \cdot \mathbf{s}$$

extending over all elements s of γ; the components $a(s)$ are arbitrary numbers in k. Such quantities are at first abstract elements forming a linear manifold (or "vector space") $\mathfrak{r}$ of h dimensions. At the same time they serve as operators[1] in $\mathfrak{R}$:

$$\mathbf{a}f = \sum_s a(s) \cdot \mathbf{s}f. \tag{2.5}$$

This "realization" suggests how to perform *multiplication*: one has

$$\mathbf{a}_1(\mathbf{a}_2 f) = \mathbf{a}f$$

if $\mathbf{a} = \mathbf{a}_1\mathbf{a}_2$ be defined by

$$a(s) = \sum_{tt'=s} a_1(t)\, a_2^t(t'). \tag{2.6}$$

The multiplication is associative:

$$(\mathbf{a}_1\mathbf{a}_2)\mathbf{a}_3 = \mathbf{a}_1(\mathbf{a}_2\mathbf{a}_3).$$

Indeed, the left side equals

$$\sum_{tt't''=s} a_1(t)\, a_2^t(t')\, a_3^{tt'}(t''),$$

whereas the right side equals

$$\sum_{tt't''=s} a_1(t)(a_2(t')a_3^{t'}(t''))^t.$$

[1] These operators are not collineations in the same sense as the operators $f \to \mathbf{s}f$.

Not even the multiplication with multiples $\beta 1$ of the unit element 1 of γ (β in k) is commutative here: while $\beta\mathbf{a}$ is to be interpreted as the quantity with the components $\beta a(s)$, $\mathbf{a}\beta$ has the components $a(s)\beta^s$.

In this section we concentrate on the *abstract group ring* ρ and shove its representation by operators (2.5) into the background after it has served its purpose of suggesting the law of multiplication (2.6).

A part $\mathfrak{p}$ of $\mathfrak{r}$ closed with respect to addition and the operation

$$\mathbf{x} \to \mathbf{x}' = \mathbf{ax} \qquad\qquad (\mathbf{x} \to \mathbf{x}' = \mathbf{xa})$$

for any $\mathbf{a}$ in ρ is called a left (right) *invariant subspace of* $\mathfrak{r}$. $\mathfrak{p}$ is a linear subspace in the sense that it contains

$$\mathbf{x} + \mathbf{y},\ \alpha\mathbf{x} \qquad\qquad (\mathbf{x} + \mathbf{y},\ \mathbf{x}\alpha)$$

along with $\mathbf{x}$, $\mathbf{y}$ whatever the number α in k. Associating

$$(a):\ \mathbf{x}' = \mathbf{ax} \tag{2.7}$$

with the quantity $\mathbf{a}$ gives rise to the regular representation (ρ) of ρ whose space is $\mathfrak{r}$ itself:

$$\mathbf{a}(\mathbf{bx}) = (\mathbf{ab})\mathbf{x}.$$

THEOREM 2A. *A left invariant subspace* $\mathfrak{p}$ *possesses an idempotent generator* $\mathbf{e}$ (*to the right*), *i.e.*, $\mathbf{xe}$ *lies in* $\mathfrak{p}$ *for every* $\mathbf{x}$, *and* $\mathbf{xe} = \mathbf{x}$ *for every* $\mathbf{x}$ *in* $\mathfrak{p}$. *The same is true for a right invariant subspace, but one must then write* $\mathbf{ex}$ *instead of* $\mathbf{xe}$.

The theorem implies that $\mathbf{e} = 1\mathbf{e}$ is in $\mathfrak{p}$ and hence $\mathbf{ee} = \mathbf{e}$.

Its proof is almost the same as given in my book, l.c., on pp. 291–292 for the customary group ring. For a left invariant subspace $\mathfrak{p}$ it runs as follows.

We construct a projection $\mathbf{x} \to \bar{\mathbf{x}}$ of $\mathfrak{r}$ onto $\mathfrak{p}$, i.e., a substitution

$$\bar{x}(s) = \sum_t d(s, t)x(t) \tag{2.8}$$

with the two desired properties that it changes every $\mathbf{x}$ into an $\bar{\mathbf{x}}$ in $\mathfrak{p}$ and is the identity within $\mathfrak{p}$. From the expression

$$y(s) = \sum_r a(r)x^r(r^{-1}s) \quad \text{for} \quad y = ax$$

one concludes that the substitution leading from

$$y(s) = x^r(r^{-1}s) \text{ to } \bar{y}(s) = \bar{x}^r(r^{-1}s)$$

is a projection as well:

$$\bar{y}(s) = \sum_t d^r(r^{-1}s, r^{-1}t)y(t).$$

Hence the same holds for the "average":

$$e(s, t) = \frac{1}{h}\sum_r d^r(r^{-1}s, r^{-1}t).$$

As it satisfies

$$e^r(r^{-1}s, r^{-1}t) = e(s, t),$$

we may write

$$e(s, t) = e^t(t^{-1}s) \text{ with } e(s) = e(s, 1),$$

and then the projection

$$\bar{x}(s) = \sum_t e(s, t)x(t)$$

can be abbreviated to $\bar{\mathbf{x}} = \mathbf{xe}$.

For a right invariant subspace a projection will be of the form

$$\bar{x}(s) = \sum_t d(s, t)x^{st^{-1}}(t)$$

rather than (2.8), because this kind of substitution changes $\mathbf{x}\alpha$ into $\bar{\mathbf{x}}\alpha$ along with $\mathbf{x} \to \bar{\mathbf{x}}$, and the averaging process is to be defined by

$$e(s, t) = \frac{1}{h} \sum_r d(sr, tr) = e(st^{-1}).$$

Incidentally right- can be reduced to left-invariance by a simple process exchanging the order of factors. With a quantity $\mathbf{a}$ we associate $\hat{\mathbf{a}}$ as defined by

$$\hat{a}(s) = \check{a}^s(s^{-1}). \tag{2.9}$$

The relationship is involutorial because (2.9) entails

$$a(s) = \check{\hat{a}}^s(s^{-1}).$$

The reader is called upon to verify the

LEMMA 2B. *If* $\mathbf{a} = \mathbf{a}_1\mathbf{a}_2$, *then* $\hat{\mathbf{a}} = \hat{\mathbf{a}}_2\hat{\mathbf{a}}_1$.

From now on only *left invariant* subspaces will be considered, and the specification "left" will be dropped. We derive from the existence of the generating idempotent these consequences (loc. cit.):

THEOREM 2C. *An invariant subspace* $\mathfrak{p}$ *containing the invariant subspace* $\mathfrak{p}_1 < \mathfrak{p}$ *can be split according to* $\mathfrak{p} = \mathfrak{p}_1 + \mathfrak{p}_2$ *into* $\mathfrak{p}_1$ *and a complementary invariant subspace* $\mathfrak{p}_2$.

Proof. $\mathbf{e}_1$ being an idempotent generator of $\mathfrak{p}_1$, we have

$$\mathbf{x} = \mathbf{xe}_1 + (\mathbf{x} - \mathbf{xe}_1) = \mathbf{x}_1 + \mathbf{x}_2$$

for every element $\mathbf{x}$ of $\mathfrak{p}$. The first summand $\mathbf{x}_1$ is in $\mathfrak{p}_1$, while the second satisfies $\mathbf{x}_2\mathbf{e}_1 = 0$. Hence $\mathfrak{p}_2 = \{\mathbf{x}_2\}$ is linearly independent of $\mathfrak{p}_1$.

A given idempotent generator $\mathbf{e}$ of $\mathfrak{p}$ splits like every $\mathbf{x}$ of $\mathfrak{p}$ into two parts

$$\mathbf{e} = \mathbf{e}_1 + \mathbf{e}_2 \tag{2.10}$$

lying in $\mathfrak{p}_1$, $\mathfrak{p}_2$ respectively. As (2.10) implies the following decomposition of any $\mathbf{x}$ in $\mathfrak{p}$:

$$\mathbf{x} = \mathbf{xe} = \mathbf{xe}_1 + \mathbf{xe}_2 = \mathbf{x}_1 + \mathbf{x}_2,$$

we have in particular

$$\mathbf{e}_1\mathbf{e}_1 = \mathbf{e}_1, \qquad \mathbf{e}_1\mathbf{e}_2 = 0,$$

$$\mathbf{e}_2\mathbf{e}_1 = 0, \qquad \mathbf{e}_2\mathbf{e}_2 = \mathbf{e}_2.$$

THEOREM 2D. *A similarity projection $\mathbf{x} \to \mathbf{x}'$ of an invariant subspace $\mathfrak{p}$ upon $\mathfrak{p}'$ is generated by aft multiplication with a quantity* $\mathbf{b}:\mathbf{x}' = \mathbf{x}\mathbf{b}$.

A correspondence $\mathbf{x} \to \mathbf{x}'$ is a similarity projection with respect to the algebra (ρ) of the operators (a), (2.7), if carrying $\mathbf{x} + \mathbf{y}$ into $\mathbf{x}' + \mathbf{y}'$ (linearity) and $\mathbf{ax}$ into $\mathbf{ax}'$. The proposition follows at once if $\mathbf{b}$ is taken as the image $\mathbf{e}'$ of the idempotent generator $\mathbf{e}$ of $\mathfrak{p}$; one then has $\mathbf{eb} = \mathbf{b}$.

Invariant subspaces which can be put into a one-to-one similarity correspondence are called *similar* or *equivalent*.

3. **Formal lemmas.** We now return to the representations (2.4), (2.5) of γ and ρ by operators in $\mathfrak{R}$. Notice that

$$\mathbf{s}(f + f') = \mathbf{s}f + \mathbf{s}f', \qquad \mathbf{s}(\alpha f) = \alpha^s \cdot \mathbf{s}f.$$

For any value $i = 1, \cdots, n$ of the index i we denote by $\mathbf{f}_i$ the quantity in ρ with the components

$$f_i(s) = \mathbf{s}f_i = \sum_k u_{ik}(s) f_k^s$$

and by $\mathbf{f}$ the column $(\mathbf{f}_1, \cdots, \mathbf{f}_n)$ whose s-component is the vector $f(s) = \mathbf{s}f$. The arguments used (l.c., §4) rest on the validity of the following formal lemmas.

LEMMA 3A.

$$\mathbf{a}\cdot\mathbf{f}_i = \sum_k \alpha_{ik}\mathbf{f}_k \tag{3.1}$$

with

$$\| \alpha_{ik} \| = \sum_r a(r) U^{-1}(r).$$

Proof. The s-component of $\mathbf{a}\cdot\mathbf{f}$ is the vector

$$g(s) = \sum_r a(r)(\mathbf{r}^{-1}\mathbf{s}f)^r.$$

Apply

$$f = \mathbf{r}(\mathbf{r}^{-1}f) = U(r)(\mathbf{r}^{-1}f)^r, \qquad (\mathbf{r}^{-1}f)^r = U^{-1}(r)f$$

to $\mathbf{s}f$ rather than f and thus verify the statement of the lemma.

LEMMA 3B. $\mathbf{f}_i\cdot\mathbf{a} = \mathbf{g}_i$ *where the vector g is defined by*

$$g = \sum_r a^r(r^{-1})\cdot\mathbf{r}f = \hat{\mathbf{a}}f.$$

In other words, if $g = \hat{\mathbf{a}}f$, then $\mathbf{g} = \mathbf{f}\cdot\mathbf{a}$.

Proof. $\mathbf{f}_i\cdot\mathbf{a} = \mathbf{x}$ is indeed given by

$$x(s) = \sum_r \mathbf{sr}f_i\cdot a^{sr}(r^{-1}) = \mathbf{s}g_i.$$

LEMMA 3C. *An equation of the kind*

$$\hat{a}(s) = \sum_{i=1}^{n} \varphi_i \cdot \mathbf{s} f_i = \varphi U(s) f^s, \tag{3.2}$$

where $\varphi = (\varphi_1, \cdots, \varphi_n)$ *is a row rather than a column of numbers (contravariant vector) entails*

$$a(s) = \sum_{i=1}^{n} f_i \cdot \mathbf{s}\varphi_i$$

when this is interpreted as meaning

$$a(s) = \varphi^s U^{-1}(s) f. \tag{3.3}$$

The linear transformation

$$\| a_{ik} \| = \| \sum_s \mathbf{s} f_i \cdot \mathbf{s}\varphi_k \|,$$

that is

$$A = \sum_s U(s) f^s \varphi^s U^{-1}(s) \tag{3.4}$$

commutes with the operators **s**.

Proof. (3.3) follows at once from (3.2) by taking (2.3) into account. g being an arbitrary vector in $\mathfrak{R}$, we compute from the explicit expression (3.4):

$$\mathbf{t}A\mathbf{t}^{-1}g = U(t)A^t(\mathbf{t}^{-1}g)^t = U(t)A^t U^{-1}(t)g,$$
$$U(t)A^t U^{-1}(t) = \sum_s U(ts) f^{ts} \varphi^{ts} U^{-1}(ts) = A.$$

4. Reciprocity between group ring and commutator algebra. The object of our investigation is the *ring* $\mathfrak{A}$ *of all linear transformations* $A = \| a_{ik} \|$:

$$f'_i = \sum_{k=1}^{n} a_{ik} f_k \tag{4.1}$$

in $\mathfrak{R}$ which *commute* with the operators **s** induced in $\mathfrak{R}$ by the elements s of γ. For each A in $\mathfrak{A}$, (4.1) thus implies

$$\mathbf{s}f'_i = \sum_k a_{ik} \cdot \mathbf{s} f_k \quad \text{or} \quad \mathbf{f}'_i = \sum_k a_{ik} \mathbf{f}_k. \tag{4.2}$$

For each A in $\mathfrak{A}$ the multiple αA lies in $\mathfrak{A}$ provided the number α is self-conjugate: $\alpha^s = \alpha$ for all group elements s. Hence $\mathfrak{A}$ is an algebra in the subfield of self-conjugate numbers rather than in k itself; nevertheless we venture to speak of $\mathfrak{A}$ as the *commutator algebra*.

The complete reciprocity between $\mathfrak{r}$ under the influence of (ρ) and $\mathfrak{R}$ under the influence of $\mathfrak{A}$ can now be established in the same manner as l.c. §4. Terms like "invariant", "irreducible", "similar" or "equivalent", when applied to $\mathfrak{r}$ or $\mathfrak{R}$ refer to the algebra (ρ) of operators $\mathbf{x} \to \mathbf{ax}$ or to the algebra $\mathfrak{A}$ respectively. The term "linear subspace" is in $\mathfrak{r}$ to be interpreted as demanding closure with

respect to addition and the ordinary multiplication $\mathbf{x} \to \alpha\mathbf{x}$ (not the modified aft multiplication $\mathbf{x} \to \mathbf{x}\alpha$) by numbers α in k. $\mathfrak{p}$ being such a linear subspace of $\mathfrak{r}$ we introduce the corresponding subspace $\mathfrak{P} = \#\mathfrak{p}$ of $\mathfrak{R}$ as *the set to which a vector f belongs if $\mathbf{f}_i$ is in $\mathfrak{p}$* for $i = 1, \cdots, n$. (4.2) shows at once that $\mathfrak{P}$ is invariant. Vice versa, if $\mathfrak{P}$ is a given linear subspace of $\mathfrak{R}$, we define $\mathfrak{p} = \natural\mathfrak{P}$ as the *linear closure of all the quantities $\mathbf{f}_i$ $(i = 1, \cdots, n)$ arising from vectors f in $\mathfrak{P}$.* If

$$f^{(\alpha)} = (f_1^{(\alpha)}, \cdots, f_n^{(\alpha)}) \qquad (\alpha = 1, 2, \cdots, m)$$

is a linear basis of the m-dimensional vector space $\mathfrak{P}$, then $\natural\mathfrak{P}$ consists of all quantities $\mathbf{x}$ of the form

$$\mathbf{x} = \sum_{\alpha, i} \varphi_i^{(\alpha)} \mathbf{f}_i^{(\alpha)} = \sum_{\alpha} (\varphi^{(\alpha)} \mathbf{f}^{(\alpha)}) \qquad (\varphi_i^{(\alpha)} \text{ arbitrary}).$$

In particular we set

$$\natural\mathfrak{R} = \mathfrak{r}_0 .$$

According to Lemma 3A, $\natural\mathfrak{P}$ is an *invariant* subspace of $\mathfrak{r}_0$. Moreover, by definition,

$$\natural\#\mathfrak{p} < \mathfrak{p}, \qquad \mathfrak{P} < \#\natural\mathfrak{P}.$$

Inclusion can here be replaced by equality: $\#$ and $\natural$ are inverse operations provided we limit ourselves within $\mathfrak{R}$ to the invariant subspaces $\mathfrak{P}$ and within $\mathfrak{r}$ to the invariant subspaces $\mathfrak{p}$ of $\mathfrak{r}_0$. Besides, those operations are conservative as to reduction, decomposition and equivalence. We exhibit these facts in two theorems:

THEOREM 4A. *If $\mathfrak{p}$ $(\mathfrak{p}', \mathfrak{p}_1, \mathfrak{p}_2)$ are any invariant subspaces of $\mathfrak{r}_0$ and $\mathfrak{P} = \#\mathfrak{p}$, then*

$$\mathfrak{p}' < \mathfrak{p}, \qquad \mathfrak{p} = \mathfrak{p}_1 + \mathfrak{p}_2, \qquad \mathfrak{p}_1 \sim \mathfrak{p}_2$$

imply

$$\mathfrak{P}' < \mathfrak{P}, \qquad \mathfrak{P} = \mathfrak{P}_1 + \mathfrak{P}_2, \qquad \mathfrak{P}_1 \sim \mathfrak{P}_2$$

respectively, while conversely,

$$\mathfrak{p} = \natural\mathfrak{P}.$$

THEOREM 4B. *If $\mathfrak{P}$ $(\mathfrak{P}', \mathfrak{P}_1, \mathfrak{P}_2)$ are any invariant subspaces of $\mathfrak{R}$ and $\mathfrak{p} = \natural\mathfrak{P}$, then*

$$\mathfrak{P} = \#\mathfrak{p}$$

and

$$\mathfrak{P}' < \mathfrak{P}, \qquad \mathfrak{P} = \mathfrak{P}_1 + \mathfrak{P}_2, \qquad \mathfrak{P}_1 \sim \mathfrak{P}_2$$

imply

$$\mathfrak{p}' < \mathfrak{p}, \qquad \mathfrak{p} = \mathfrak{p}_1 + \mathfrak{p}_2, \qquad \mathfrak{p}_1 \sim \mathfrak{p}_2$$

respectively.

Before proceeding to the almost literal repetition of the proofs, we make this remark. If $\hat{\mathbf{e}}$ is the idempotent generator of an invariant $\mathfrak{p}$, then the corresponding $\mathfrak{P} = \#\mathfrak{p}$ consists of all vectors of the form $\mathbf{e}f$. Indeed, $g = \mathbf{e}f$ lies in $\mathfrak{P}$ because $\mathbf{g}_i = \mathbf{f}_i\hat{\mathbf{e}}$ by Lemma 3B, and for each f in $\mathfrak{P}$ one has $\mathbf{e}f = f$.

1. We prove the first part of Theorem 4A by observing that the decomposition $\mathfrak{p} = \mathfrak{p}_1 + \mathfrak{p}_2$ when applied to an idempotent generator $\hat{\mathbf{e}}$ of $\mathfrak{p}$: $\hat{\mathbf{e}} = \hat{\mathbf{e}}_1 + \hat{\mathbf{e}}_2$ leads to this decomposition $\mathfrak{P} = \mathfrak{P}_1 + \mathfrak{P}_2$ of $\mathfrak{P} = \#\mathfrak{p}$:

$$F = \mathbf{e}F = \mathbf{e}_1F + \mathbf{e}_2F = F_1 + F_2 \qquad (F \text{ in } \mathfrak{P}).$$

Lemma 2B allows us to shear all the roofs off the relations

$$\hat{e}_1\hat{e}_2 = \hat{e}_2\hat{e}_1 = 0 \qquad (\hat{e}_1\hat{e}_1 = \hat{e}_1,\ \hat{e}_2\hat{e}_2 = \hat{e}_2),$$

thus warranting the independence of the parts $\mathfrak{P}_1$, $\mathfrak{P}_2$:

$$\mathbf{e}_1F_2 = 0, \qquad \mathbf{e}_2F_1 = 0.$$

2. The similarity correspondence between $\mathfrak{p}_1$ and $\mathfrak{p}_2$,

$$\mathbf{x}_2 = \mathbf{x}_1\hat{\mathbf{b}}, \qquad \mathbf{x}_1 = \mathbf{x}_2\hat{\mathbf{b}}',$$

gives rise to the mutually inverse transformations

$$f_2 = \mathbf{b}f_1, \qquad f_1 = \mathbf{b}'f_2$$

between the vectors f_1, f_2 of $\mathfrak{P}_1 = \#\mathfrak{p}_1$ and $\mathfrak{P}_2 = \#\mathfrak{p}_2$. By (4.2) these formulas establish a similarity correspondence:

$$f_1' = Af_1 \text{ entails } \mathbf{b}f_1' = A(\mathbf{b}f_1) \qquad \{A \text{ in } \mathfrak{A}\}.$$

To secure the last part, $\mathfrak{p} < \natural\mathfrak{P}$, we construct $\natural\mathfrak{P}$ by means of the idempotent generator $\hat{\mathbf{e}}$ of $\mathfrak{p}$ as follows: if $g^{(\alpha)}$ $(\alpha = 1, \cdots, n)$ ranges over a basis of the complete vector space $\mathfrak{R}$, all $f^{(\alpha)} = \mathbf{e}g^{(\alpha)}$ lie in $\mathfrak{P} = \#\mathfrak{p}$, and hence

$$\mathbf{y} = \sum_{\alpha,i} \varphi_i^{(\alpha)}\mathbf{f}_i^{(\alpha)} = \sum_{\alpha} (\varphi^{(\alpha)}\mathbf{f}^{(\alpha)})$$

in $\natural\mathfrak{P}$. On introducing

$$\mathbf{x} = \sum_{\alpha} (\varphi^{(\alpha)}\mathbf{g}^{(\alpha)}),$$

we have $\mathbf{y} = \mathbf{x}\hat{\mathbf{e}}$. So $\mathbf{x}\hat{\mathbf{e}}$ lies in $\natural\mathfrak{P}$ if $\mathbf{x}$ lies in $\mathfrak{r}_0 = \natural\mathfrak{R}$. But each $\mathbf{x}$ in $\mathfrak{p}$ satisfies both conditions: $\mathbf{x}$ in $\mathfrak{r}_0$ and $\mathbf{x}\hat{\mathbf{e}} = \mathbf{x}$.

The converse Theorem 4B exhibits the really important facts. Its assertion that $\mathfrak{p} = \natural\mathfrak{P}$ implies $\mathfrak{P} = \#\mathfrak{p}$ for any subspace $\mathfrak{P}$ invariant under $\mathfrak{A}$ is the backbone of the whole theory. Let $\hat{\mathbf{e}}$ be the idempotent generator of $\mathfrak{p} = \natural\mathfrak{P}$. Like all elements of $\mathfrak{p}$ it is of the form

$$\hat{e}(s) = \sum_{\alpha,i} \varphi_i^{(\alpha)} \cdot \mathbf{s}f_i^{(\alpha)},$$

where

$$f^{(\alpha)} = (f_1^{(\alpha)}, \cdots, f_n^{(\alpha)})$$

ranges over a basis of $\mathfrak{P}$. Hence by Lemma 3C,

$$e(s) = \sum_{\alpha,k} \mathbf{s}\varphi_k^{(\alpha)} \cdot f_k^{(\alpha)},$$

and any vector $g = \mathbf{e}f$ of $\#\mathfrak{p}$ is given by

$$g_i = \sum_\alpha \Big(\sum_k a_{ik}^{(\alpha)} f_k^{(\alpha)}\Big) = \sum_\alpha g_i^{(\alpha)}, \tag{4.3}$$

where

$$a_{ik}^{(\alpha)} = \sum_s \mathbf{s}f_i \cdot \mathbf{s}\varphi_k^{(\alpha)}.$$

Each term $g^{(\alpha)}$ of the sum (4.3), $g = \Sigma g^{(\alpha)}$, arises from $f^{(\alpha)}$ by a linear transformation $A^{(\alpha)} = \| a_{ik}^{(\alpha)} \|$ which according to the same lemma commutes with all $\mathbf{s}$. Hence $\mathfrak{P}$, being invariant with respect to the transformations A of the commutator algebra $\mathfrak{A}$, contains $g^{(\alpha)}$ as well as $f^{(\alpha)}$. This proves our statement: g in $\mathfrak{P}$ or $\#\mathfrak{p} < \mathfrak{P}$.

The decomposition $\mathfrak{P} = \mathfrak{P}_1 + \mathfrak{P}_2$ implies by definition that each quantity $\mathbf{x}$ in $\mathfrak{p} = \natural\mathfrak{P}$ can be written as a sum $\mathbf{x}_1 + \mathbf{x}_2$, $\mathbf{x}_1$ in $\mathfrak{p}_1 = \natural\mathfrak{P}_1$, $\mathbf{x}_2$ in $\mathfrak{p}_2 = \natural\mathfrak{P}_2$. It remains to prove that $\mathfrak{p}_1$ and $\mathfrak{p}_2$ are linearly independent or that the intersection $\mathfrak{p}^* = \mathfrak{p}_1 \circ \mathfrak{p}_2$ is empty provided $\mathfrak{P}^* = \mathfrak{P}_1 \circ \mathfrak{P}_2$ be empty. But according to the part of Theorem 4B already proved,

$$\#\mathfrak{p}^* < \#\mathfrak{p}_1 = \mathfrak{P}_1, \qquad \#\mathfrak{p}^* < \mathfrak{P}_2;$$

hence $\#\mathfrak{p}^* < \mathfrak{P}^*$ and by the last part of Theorem 4A: $\mathfrak{p}^* < \natural\mathfrak{P}^*$.

The transition from $\mathfrak{P}_1 \sim \mathfrak{P}_2$ to $\mathfrak{p}_1 \sim \mathfrak{p}_2$ for $\mathfrak{p}_1 = \natural\mathfrak{P}_1$, $\mathfrak{p}_2 = \natural\mathfrak{P}_2$, is to be based on the following statement, the proof of which is contained in Lemma 3B:

Lemma 4C. *$\mathfrak{r}_0$ is right as well as left invariant.*

Therefore $\mathfrak{r}_0$ has an idempotent generator $\mathbf{i}$ to the left: $\mathbf{i}$ in $\mathfrak{r}_0$, $\mathbf{ix} = \mathbf{x}$ for every $\mathbf{x}$ in $\mathfrak{r}_0$.

Let $f^{(\alpha)}$ be a basis for $\mathfrak{P}_1$ and let the given similar mapping of $\mathfrak{P}_1$ on $\mathfrak{P}_2$ send $f^{(\alpha)}$ into $g^{(\alpha)}$. When we put

$$\mathbf{x} = \sum_{\alpha,i} \varphi_i^{(\alpha)} \cdot \mathbf{f}_i^{(\alpha)}, \qquad \mathbf{y} = \sum_{\alpha,i} \varphi_i^{(\alpha)} \cdot \mathbf{g}_i^{(\alpha)}, \qquad (\varphi_i^{(\alpha)} \text{ arbitrary numbers})$$

the correspondence $\mathbf{x} \to \mathbf{y}$ between an $\mathbf{x}$ and a $\mathbf{y}$ with the same coefficients $\varphi_i^{(\alpha)}$ will establish a similarity mapping of $\mathfrak{p}_1 = \natural\mathfrak{P}$ on $\mathfrak{p}_2 = \natural\mathfrak{P}_2$, because by Lemma 3A, we obtain

$$\mathbf{ax} = \sum_{\alpha,k} \psi_k^{(\alpha)} \mathbf{f}_k^{(\alpha)}, \qquad \mathbf{ay} = \sum_{\alpha,k} \psi_k^{(\alpha)} \mathbf{g}_k^{(\alpha)}$$

with

$$\psi_k^{(\alpha)} = \sum_i \varphi_i^{(\alpha)} \cdot \alpha_{ik}.$$

This definition of $\mathbf{x} \to \mathbf{y}$, however, goes through only *if* $\mathbf{x} = 0$ *implies* $\mathbf{y} = 0$. We first prove that

$$\hat{\mathbf{x}}f = 0 \text{ implies } \hat{\mathbf{y}}f = 0$$

(for any vector f). By Lemma 3C the vector $F = \hat{\mathbf{x}}f$ is a sum of terms $F^{(\alpha)}$, the α-th of which arises from $f^{(\alpha)}$ by the transformation

$$A^{(\alpha)} = \| a_{ik}^{(\alpha)} \| = \left\| \sum_s \mathbf{s}f_i \cdot \mathbf{s}\varphi_k^{(\alpha)} \right\|.$$

Since $A^{(\alpha)}$ is in $\mathfrak{A}$, the given similarity mapping of $\mathfrak{P}_1$ on $\mathfrak{P}_2$ sends $F^{(\alpha)}$ into the corresponding part $G^{(\alpha)}$ of $G = \hat{\mathbf{y}}f$ and hence F into G. Therefore $F = 0$ implies $G = 0$, and more especially, when the numbers $\varphi_i^{(\alpha)}$ satisfy the equation $\mathbf{x} = 0$, we must have $\hat{\mathbf{y}}f = 0$ for every vector f, or by Lemma 3B, $\mathbf{f}_i \cdot \mathbf{y} = 0$. Hence the given quantity $\mathbf{y}$ satisfies $\mathbf{zy} = 0$ for every $\mathbf{z}$ in $\mathfrak{r}_0$, in particular for $\mathbf{z} = \mathbf{i}$. But as $\mathbf{y}$ itself lies in $\mathfrak{r}_0$, the ensuing equation $\mathbf{iy} = 0$ yields the desired result: $\mathbf{y} = 0$.

The complete reciprocity established by Theorems 4A and B involves the fact that the process $\#$ not only changes $\mathfrak{p} = 0$ into $\#\mathfrak{p} = 0$ and a part $\mathfrak{p}' < \mathfrak{p}$ into a part $\#\mathfrak{p}' < \#\mathfrak{p}$, but also a $\mathfrak{p} \neq 0$ into a $\#\mathfrak{p} \neq 0$ and a *proper* part into a *proper* part *provided the* $\mathfrak{p}$'s *are invariant subspaces of* $\mathfrak{r}_0$. *The decomposition of* $\mathfrak{r}_0$ *into irreducibly invariant subspaces* $\mathfrak{p}$ *leads to a decomposition of* $\mathfrak{R}$ *into subspaces* $\mathfrak{P}$ *irreducibly invariant under the algebra* $\mathfrak{A}$, *and both decompositions run absolutely parallel even as to the pertaining equivalences.*

5. **Representations of the group ring. The roof operation.** Not so simple is the relationship of the abstract group ring of our quantities $\mathbf{a}$ to its representation by the operators

$$f \to \mathbf{a}f$$

in $\mathfrak{R}$ which form a *homomorphic operator algebra* $\mathfrak{B}$, although $\mathfrak{B}$ also breaks up into irreducible parts similar to the irreducible parts of ρ. For a given vector f and a given invariant subspace $\mathfrak{p}$ of $\mathfrak{r}$ we denote by $\mathfrak{p}(f)$ the set of vectors $\mathbf{x}f$ arising from f by all the $\mathbf{x}$ in $\mathfrak{p}$. Let $f^{(\alpha)}$ $(\alpha = 1, \cdots, n)$ be a basis for the n-space $\mathfrak{R}$ and

$$\mathfrak{r} = \mathfrak{p}_1 + \mathfrak{p}_2 + \cdots \tag{5.1}$$

a decomposition of $\mathfrak{r}$ into subspaces irreducibly invariant under (ρ). Our statement is proved in the well-known manner by going through the list

$$\begin{array}{l} \mathfrak{p}_1(f^{(1)}), \cdots, \mathfrak{p}_1(f^{(n)}), \\ \mathfrak{p}_2(f^{(1)}), \cdots, \mathfrak{p}_2(f^{(n)}), \\ \cdots\cdots\cdots\cdots \end{array}$$

and dropping a term each time it is contained in the sum of the preceding ones. The construction depends on the choice of the coördinate system $f^{(\alpha)}$ as well as

the decomposition (5.1), and does not result in such a thoroughgoing parallelism as we encountered for the commutator group. Coincidence between the numbers of equivalent parts is not to be expected. Whereas for the commutator algebra it was essential to restrict oneself to the two-sided invariant subspace $\mathfrak{r}_0$, $\mathfrak{r}$ is here to be taken modulo that two-sided invariant subspace whose elements $\mathbf{a}$ satisfy the equation $\mathbf{a}f = 0$ identically in the vector f.

Considering all these circumstances, it seems to me inadequate to avail oneself of the reciprocity[2] between a matric algebra $\mathfrak{B}$ and its commutator algebra $\mathfrak{A}$ for getting a hold on $\mathfrak{A}$ through $\mathfrak{B}$, although the most essential point, the full reducibility of $\mathfrak{A}$, can be reached by this method applicable to any abstract semi-simple algebra ρ. One could visualize the situation as follows. One decomposes the "one" 1 of ρ into independent primitive idempotents:

$$1 = \mathbf{e}_1 + \mathbf{e}_2 + \cdots;$$

$\mathbf{e}_i\mathbf{e}_k = \mathbf{e}_i$ or 0 according as $i = k$ or $i \neq k$. (An idempotent $\mathbf{e}$ is *primitive* if not allowing of a decomposition $\mathbf{e}_1 + \mathbf{e}_2$ into two independent idempotents except the trivial ones $\mathbf{e} + 0$ and $0 + \mathbf{e}$.) It can be shown then that

$$\mathbf{x} = \mathbf{x}\mathbf{e}_1 + \mathbf{x}\mathbf{e}_2 + \cdots \text{ and } f = \mathbf{e}_1 f + \mathbf{e}_2 f + \cdots$$

result in a decomposition of $\mathfrak{r}$ and $\mathfrak{R}$ respectively,

$$\mathfrak{r} = \mathfrak{p}_1 + \mathfrak{p}_2 + \cdots, \qquad \mathfrak{R} = \mathfrak{P}_1 + \mathfrak{P}_2 + \cdots,$$

into irreducibly invariant subspaces. Again, $f \to \mathbf{a}f$ is a given representation of the elements $\mathbf{a}$ of ρ by operators in $\mathfrak{R}$, and invariance in $\mathfrak{r}$ refers to the operations $\mathbf{x} \to \mathbf{ax}$ in $\mathfrak{R}$ to the algebra $\mathfrak{A}$ of linear transformations commuting with the operations $f \to \mathbf{a}f$. However, the correspondence thus established between the parts $\mathfrak{p}_i$ of $\mathfrak{r}$ and $\mathfrak{P}_i$ of $\mathfrak{R}$ depends on the choice of the idempotent generators $\mathbf{e}_i$ of $\mathfrak{p}_i$. To make the correspondence independent, one must match the subspace $\mathfrak{P}$ of the vectors $\mathbf{e}f$ against the subspace $\mathfrak{p}$ of the quantities $\mathbf{x}\hat{\mathbf{e}}$ rather than $\mathbf{xe}$. So one may say that the more elementary and complete reciprocity we expatiated on in the preceding sections is due to the existence of the operation $\hat{}$ in a group ring while missing in an arbitrary abstract algebra. The rôle of $\hat{}$ is further clarified by the following.[3]

THEOREM 5A. *If the invariant subspaces $\mathfrak{p}_1$, $\mathfrak{p}_2$ generated by the idempotents e_1, e_2 are equivalent to each other, so are the invariant subspaces $\hat{\mathfrak{p}}_1$, $\hat{\mathfrak{p}}_2$ generated by $\hat{e}_1$, $\hat{e}_2$. $\mathfrak{p}$ and $\hat{\mathfrak{p}}$ are the substrata of contragredient representations.*

Let the one-to-one similarity mapping $x_1 \to x_2$ of $\mathfrak{p}_1$ on $\mathfrak{p}_2$ carry e_1 into b and in the inverse direction e_2 into a. We then have

$$x_2 = x_1 b, \qquad x_1 = x_2 a. \tag{5.2}$$

[2] See, for instance, Weyl, Ann. Math., vol. 37 (1936), p. 718, Th. (1.4-B). Even the above statement should be made with reservation, since $\mathfrak{B}$ is not a *matric* algebra, and $\mathfrak{A}$ not a matric *algebra* in the strict sense.

[3] Compare (l. c.) pp. 352–354.

b satisfies $e_1 b = b$, and like every other element of $\mathfrak{p}_2$, $be_2 = b$. Hence

$$(5.3) \qquad e_1 b e_2 = b, \qquad e_2 a e_1 = a.$$

Moreover, if we put $x_1 = a$ in the first and $x_2 = b$ in the second of the equations (5.2), we obtain

$$(5.4) \qquad e_2 = ab, \qquad e_1 = ba.$$

Conversely the relations (5.4) guarantee that (5.2) are reciprocal mappings $\mathfrak{p}_1 \rightleftarrows \mathfrak{p}_2$. We need only to "roof" these equations (5.3), (5.4) in order to conclude that $\hat{e}_1$, $\hat{e}_2$ are linked by $\hat{b}$, $\hat{a}$ in the same fashion as e_1, e_2 by a, b.

To prove the second part we have to introduce the notion of *trace*: the trace of a quantity a, $\operatorname{tr}(a)$, is its unit component $a(1)$. The trace $\operatorname{tr}(xy)$ of the product of two variable quantities x, y is bilinear in the sense that it satisfies the distributive law with respect to decomposition $x = x_1 + x_2$ of the first as well as the second factor, and that it takes on the numerical factor α if one replaces x by αx or y by $y\alpha$ (distinguish between fore and aft multiplication). Moreover,

$$(5.5) \qquad \operatorname{tr}(xy) = \sum_s x(s) y^s(s^{-1})$$

is a *non-degenerate* bilinear form: each of the equations $\operatorname{tr}(ax) = 0$ or $\operatorname{tr}(xa) = 0$ when holding identically in x leads to $a = 0$. (5.5) is not symmetric in x and y as is the case for the ordinary group ring. However, the obvious equation $\operatorname{tr}(\hat{a}) = \operatorname{tr}(a)$, together with Lemma 2B, establish the following modified law of symmetry:

$$\operatorname{tr}(yx) = \operatorname{tr}(\hat{x}\hat{y}).$$

One readily verifies

$$(5.6) \qquad \operatorname{tr}(sxs^{-1}) = \operatorname{tr}^s(x).$$

e being a given idempotent, let $\mathfrak{p}$ and $\mathfrak{q}$ be the left and the right invariant subspaces consisting of the quantities xe and ex respectively. We assert that $\operatorname{tr}(xy)$ is non-degenerate if x and y vary in $\mathfrak{p}$ and $\mathfrak{q}$ respectively. Indeed, if z is any element whatever and a in $\mathfrak{p}$, then

$$az = ae \cdot z = a \cdot ez = ay,$$

where $y = ez$ is in $\mathfrak{q}$. Hence the assumption $\operatorname{tr}(ay) = 0$ for y in $\mathfrak{q}$ implies $\operatorname{tr}(az) = 0$ for all z, whence $a = 0$. Similarly for the second factor. We now refer $\mathfrak{p}$ and $\mathfrak{q}$ each to a coördinate system a_i and b_k such that

$$(5.7) \qquad x = \xi_1 a_1 + \cdots + \xi_g a_g, \qquad y = b_1 \eta_1 + \cdots + b_h \eta_h$$

describe $\mathfrak{p}$ and $\mathfrak{q}$ if the numbers ξ and η vary freely in k. From the non-degeneracy of $\operatorname{tr}(xy)$ for (5.7) follows readily the coincidence $h = g$ of the dimensions of $\mathfrak{p}$ and $\mathfrak{q}$ and the possibility of adapting the coördinate system b_k in $\mathfrak{q}$ to the arbitrarily chosen coördinate system a_i in $\mathfrak{p}$ such that

$$\operatorname{tr}(xy) = \xi_1 \eta_1 + \cdots + \xi_g \eta_g$$

for x in $\mathfrak{p}$ and y in $\mathfrak{q}$.

We now consider the simultaneous substitutions

(5.8) $$x' = sx, \qquad y' = ys^{-1}$$

in $\mathfrak{p}$ and $\mathfrak{q}$ respectively. The y-substitution may also be put into the form

(5.9) $$\hat{y}' = s\hat{y},$$

where

$$\hat{y} = \eta_1\hat{b}_1 + \cdots + \eta_g\hat{b}_g$$

now varies in the *left* invariant subspace $\hat{\mathfrak{p}}$ generated by $\hat{e}$. We write (5.8), (5.9) in terms of the coördinates as

$$\xi'_i = \sum_k u_{ik}(s)\xi^s_k, \qquad \eta'_i = \sum_k v_{ik}(s)\eta^s_k.$$

The relation (5.6) yielding

$$\mathrm{tr}(x'y') = \mathrm{tr}^s(xy) \quad \text{or} \quad \sum_i \xi'_i\eta'_i = \sum_i \xi^s_i\eta^s_i$$

proves the two matrices

$$\| u_{ik}(s) \|, \qquad \| v_{ik}(s) \|$$

to be *contragredient*. Hence the regular representation of ρ induces in $\mathfrak{p}$ and $\hat{\mathfrak{p}}$ contragredient representations of γ (assuming coördinate systems in $\mathfrak{p}$ and $\hat{\mathfrak{p}}$ properly adapted to each other). Observe that passage to the contragredient matrices $\hat{U}(s)$ in a collinear representation $s \to U(s)$ produces a *representation* again, as is readily seen from equation (2.2).

THE INSTITUTE FOR ADVANCED STUDY.

TAYLOR'S SERIES OF ENTIRE FUNCTIONS OF SMOOTH GROWTH

By N. Wiener and W. T. Martin

1. **Introduction.**[1] Let $a_n \geqq 0$ $(n = 1, 2, \cdots)$, and

$$\sum_1^\infty a_n x^n \sim sH(x) \qquad (x \to \infty). \tag{1}$$

We may rewrite (1) in the form

$$\sum_1^\infty a_n e^{n\xi - F(\xi)} \to s \qquad (\xi \to \infty), \tag{2}$$

where

$$H(x) = e^{F(\log x)}. \tag{3}$$

We shall derive the following Tauberian theorem.

Theorem 1. *Let $a_n \geqq 0$ $(n = 1, 2, \cdots)$, and let* (2) *hold, where*

(4) *F is four times continuously differentiable for $a \leqq x$ for some a;*

$$F''(x) \geqq \text{const.} > 0 \textit{ for } a \leqq x; \tag{5}$$

$$F'''(x) = o([F''(x)]^{3/2}) \qquad (x \to \infty); \tag{6}$$

$$F^{\text{iv}}(x) = o([F''(x)]^2) \qquad (x \to \infty). \tag{7}$$

Then for any positive λ,

$$\lim_{x\to\infty} \frac{(2\pi)^{1/2}}{\lambda} \sum_{x \leqq \psi(n) \leqq x+\lambda} a_n e^{nG(n) - F(G(n))} = s, \tag{8}$$

where $\psi(x)$ is defined by

$$\int_a^x [G'(x)]^{1/2} dx = \psi(x), \tag{9}$$

and $G(x)$ is the inverse function to $F'(x)$ for $a \leqq x$.

As a converse to this theorem we shall prove

Theorem 2. *Let $a_n \geqq 0$ $(n = 1, 2, \cdots)$, and let* (8) *hold for every positive value of λ, where F fulfills the conditions* (4), (5), (6) *and* (7) *and ψ and G are defined as in Theorem* 1. *Then* (2) *holds.*

Received January 19, 1937.

[1] The authors' attention was directed to problems of this type by Professor Vijayaraghavan. It has come to our attention that a similar group of problems has been recently attacked by Mr. Kales of Brown University by quite different methods. While neither direction of work is reducible to the other, Mr. Kales' priority in entering the field is clear.

In the next theorem we place some additional restrictions on F and we obtain certain inequalities related to ψ.

THEOREM 3. (i) *Let F fulfill the conditions* (4) *and* (5) *and the further condition*

$$\text{(10)} \quad \textit{for some positive } \epsilon,\ F''(x) = O([F'(x)]^{1+\epsilon}) \qquad (x \to \infty).$$

Then for any positive ϵ,

$$\text{(11)} \qquad n^{-\frac{1}{2}-\epsilon} < \psi'(n)$$

and

$$\text{(12)} \qquad n^{\frac{1}{2}-\epsilon} < \psi(n),$$

for sufficiently large values of n. (ii) *Let F fulfill the conditions* (4) *and* (5) *and the additional condition*

(13) *$F'(x)$ is ultimately greater than any power of x.*

Then for any positive ϵ,

$$\text{(14)} \qquad \psi(n) < n^{\frac{1}{2}+\epsilon}$$

for sufficiently large values of n.

We shall obtain the following theorems on entire functions of smooth growth as immediate consequences of these Tauberian theorems.

THEOREM 4. *Let*

$$\text{(15)} \qquad f(x) = \sum_0^\infty b_n x^n$$

be an integral function, and let

$$\text{(16)} \qquad \frac{1}{2\pi}\int_0^{2\pi} |f(re^{i\theta})|^2 d\theta \sim se^{F(2\log r)} \qquad (r \to \infty),$$

where F fulfills the conditions (4), (5), (6) *and* (7). *Then for any positive* λ,

$$\text{(17)} \qquad \lim_{x\to\infty} \frac{(2\pi)^{\frac{1}{2}}}{\lambda} \sum_{x \leq \psi(n) \leq x+\lambda} |b_n|^2 e^{nG(n)-F(G(n))} = s,$$

where ψ and G are defined as in Theorem 1.

The converse to Theorem 4 is

THEOREM 5. *Let* (15) *be an entire function and let* (17) *hold for every positive value of λ, where F fulfills the conditions* (4), (5), (6) *and* (7), *and G and ψ are defined as in Theorem* 1. *Then* (16) *holds.*

If we place additional conditions on F, we have the following gap theorem.

THEOREM 6. *Let* (15) *be an entire function, and let* (16) *hold, where $s \neq 0$ and F satisfies the conditions* (4), (5), (6), (7), (10) *and* (13). *Then* (17) *holds where ψ and G are defined as in Theorem* 1. *Furthermore, for any positive ϵ we shall always have* (11) *and*

$$\text{(18)} \qquad n^{\frac{1}{2}-\epsilon} < \psi(n) < n^{\frac{1}{2}+\epsilon}$$

for sufficiently large values of n. Thus the function (15) *can have only a finite number of gaps of magnitude* $(\nu, \nu + \nu^{\frac{1}{2}+\epsilon})$, *that is, for any positive* ϵ, *the equations*

$$a_n = a_{n+1} = \cdots = a_{n+[n^{\frac{1}{2}+\epsilon}]} = 0$$

can hold for at most a finite number of values of n.

As a consequence of Theorems 1 and 2 we shall derive the following non-linear Tauberian theorem.

THEOREM 7. *Let* $a_n \geqq 0$ $(n = 1, 2, \cdots)$, *and let*

$$c_n = \sum_1^n a_m a_{n-m} \qquad (n = 1, 2, \cdots). \tag{19}$$

If F fulfills the conditions (4), (5), (6) *and* (7) *and if ψ and G are defined as in Theorem* 1, *then the following two statements are equivalent*: (i) *the relation* (8) *holds for every positive value of* λ; (ii) *the relation*

$$\lim_{x\to\infty} \frac{(2\pi)^{\frac{1}{2}}}{\lambda} \sum_{x \leqq \sqrt{2}\psi(\frac{1}{2}n) \leqq x+\lambda} c_n e^{nG(\frac{1}{2}n) - 2F(G(\frac{1}{2}n))} = s^2 \tag{20}$$

holds for every positive value of λ.

For the proof of Theorem 1 we shall put

$$\xi = \varphi(u + \psi(n)), \tag{21}$$

so that u occurs as the difference of two variables

$$u = \varphi^{-1}(\xi) - \psi(n). \tag{22}$$

Here ϕ and ψ are functions to be determined with regard to the following considerations. We shall develop the exponent $n\xi - F(\xi)$ in a Taylor's series with remainder and we shall determine ϕ and ψ in such a manner that the coefficient of u shall vanish and that the coefficient of $\frac{1}{2}u^2$ shall be negative unity. Then we shall use the conditions on F to show that the remainder in the Taylor's series approaches zero as $\xi \to \infty$; indeed, that it approaches zero in such a manner that (2) shall imply that

$$\sum_1^\infty a_n e^{nG(n) - F(G(n))} e^{-\frac{1}{2}u^2} \to s \qquad (\xi \to \infty). \tag{23}$$

We shall then show that the General Tauberian Theorem of Wiener[2] is applicable to (23). Here the kernel $K(u)$ is $e^{-\frac{1}{2}u^2}$ whose Fourier transform $e^{-\frac{1}{2}x^2}$ is non-vanishing, and, as we have noted, u occurs as the difference of two variables. The relation (8) will follow immediately.

Theorem 2, the converse of Theorem 1, will follow by an application of the converse portion of the General Tauberian Theorem just referred to. The other theorems will follow at once.

2. Let us put

$$\xi = \varphi(u + \psi(n)). \tag{21}$$

[2] N. Wiener, *Tauberian theorems*, Annals of Mathematics, (2), vol. 33 (1932), pp. 1–100.

We have, formally, on holding n fast and developing in powers of u,

$$
\begin{aligned}
(24)\quad n\xi - F(\xi) &= n\varphi(u + \psi(n)) - F(\varphi(u + \psi(n))) \\
&= n\varphi(\psi(n)) - F(\varphi(\psi(n))) \\
&+ u[n - F'(\varphi(\psi(n)))]\varphi'(\psi(n)) \\
&+ \frac{u^2}{2}\{[n - F'(\varphi(\psi(n)))]\varphi''(\psi(n)) - F''(\varphi(\psi(n)))[\varphi'(\psi(n))]^2\} \\
&+ \frac{u^3}{6} D_v^3[n\varphi(v) - F(\varphi(v))]_{v=\theta u + \psi(n)},
\end{aligned}
$$

where $0 \leqq \theta = \theta(u, n) \leqq 1$.

We desire that the coefficient of u shall be zero, and the coefficient of $\frac{1}{2}u^2$ negative unity, whether or not n is an integer. Necessary conditions for this are, formally, that

$$(25)\qquad F'(\varphi(\psi(w))) = w,$$

$$(26)\qquad \varphi'(\psi(w)) = \psi'(w),$$

for, as regards the latter, the coefficient of $\frac{1}{2}u^2$ is

$$(27)\qquad \{D_w[w\varphi'(\psi(w)) - F'(\varphi(\psi(w)))\varphi'(\psi(w))] - \varphi'(\psi(w))\}/\psi'(w).$$

Denoting by $G(w)$ the inverse function to $F'(w)$, we see that (25) and (26) become

$$(28)\qquad \varphi(\psi(w)) = G(w),$$

$$(29)\qquad [\psi'(w)]^2 = G'(w).$$

For our functions φ, ψ we are thus led to take

$$(30)\qquad \psi(w) = \int_a^w [G'(t)]^{\frac{1}{2}}\,dt, \qquad \varphi(w) = G(\psi^{-1}(w)),$$

where a is so large that F is four times continuously differentiable for $w \geqq a$ and $0 <$ const. $\leqq F''(w)$. With ψ, φ so defined, (24), with w in place of n, is valid for w and $u + \psi(w)$ sufficiently large and (25), (26), (28) and (29) are valid for $w \geqq a$.

The last term of (24) may be written

$$
\begin{aligned}
(31)\quad \frac{u^3}{6}\{&w\varphi'''(\theta u + \psi(w)) - F'''(\varphi(\theta u + \psi(w)))[\varphi'(\theta u + \psi(w))]^3 \\
&- 3F''(\varphi(\theta u + \psi(w)))\varphi'(\theta u + \psi(w))\varphi''(\theta u + \psi(w)) \\
&- F'(\varphi(\theta u + \psi(w)))\varphi'''(\theta u + \psi(w))\}.
\end{aligned}
$$

Differentiating (25) and using (26), we have

$$F''(\varphi(\psi(w)))[\varphi'(\psi(w))]^2 = 1. \tag{32}$$

For a general argument z, (32) becomes

$$F''(\varphi(z))[\varphi'(z)]^2 = 1, \tag{33}$$

and yields on differentiation

$$F'''(\varphi(z))[\varphi'(z)]^3 + 2F''(\varphi(z))\varphi'(z)\varphi''(z) = 0. \tag{34}$$

Thus if $\theta u + \psi(w) = z$, (31) becomes

$$\begin{aligned} &\frac{u^3}{6}\{w\varphi'''(z) - F''(\varphi(z))\varphi'(z)\varphi''(z) - F'(\varphi(z))\varphi'''(z)\} \\ &\qquad = \frac{u^3}{6}\left\{[F'(\varphi(\psi(w))) - F'(\varphi(z))]\varphi'''(z) - \frac{\varphi''(z)}{\varphi'(z)}\right\}. \end{aligned} \tag{35}$$

Now, by the law of the mean, there will be some v in the interval between $\psi(w)$ and z such that

$$F'(\varphi(\psi(w))) - F'(\varphi(z)) = -\theta u F''(\varphi(v))\varphi'(v) = -\frac{\theta u}{\varphi'(v)}. \tag{36}$$

Expression (35) thus becomes

$$-\frac{u^3}{6}\left\{\theta u\frac{\varphi'''(z)}{\varphi'(v)} + \frac{\varphi''(z)}{\varphi'(z)}\right\}, \tag{37}$$

and

$$e^{w\xi - F(\xi)} = \exp\left\{w\varphi(\psi(w)) - F(\varphi(\psi(w))) - \tfrac{1}{2}u^2 - \frac{u^3}{6}\left[\theta u\frac{\varphi'''(z)}{\varphi'(v)} + \frac{\varphi''(z)}{\varphi'(z)}\right]\right\}. \tag{38}$$

We shall now prove two facts relating to the function

$$A(u, \xi) = w\xi - F(\xi) - wG(w) + F(G(w)), \tag{39}$$

where

$$\xi = \varphi(u + \psi(w)). \tag{40}$$

(A) *For any positive* U

$$\max_{|u| \leqq U} |A(u, \xi) + \tfrac{1}{2}u^2| \to 0 \qquad (\xi \to \infty); \tag{41}$$

(B) *there are constants* w_0 *and* ξ_0 *such that, if*

$$2 \leqq |u|, \qquad u \leqq \varphi^{-1}(\xi) - \psi(w_0), \qquad \xi_0 \leqq \xi, \tag{42}$$

then

$$A(u, \xi) \leqq -|u|/2. \tag{43}$$

For the proof of (A) let us differentiate

$$\varphi'(\psi(x)) = [G'(x)]^{\frac{1}{2}} \tag{44}$$

to obtain

$$\varphi''(\psi(x))\psi'(x) = \tfrac{1}{2}[G'(x)]^{-\frac{1}{2}}G''(x). \tag{45}$$

Thus

$$\varphi''(\psi(x)) = \frac{G''(x)}{2G'(x)}; \tag{46}$$

$$\varphi'''(\psi(x))\psi'(x) = \frac{G'(x)G'''(x) - [G''(x)]^2}{2[G'(x)]^2}; \tag{47}$$

$$\frac{\varphi''(\psi(x))}{\varphi'(\psi(x))} = \frac{G''(x)}{2[G'(x)]^{\frac{3}{2}}}; \tag{48}$$

$$\frac{\varphi'''(\psi(x))}{\varphi'(\psi(x))} = \frac{G'(x)G'''(x) - [G''(x)]^2}{2[G'(x)]^3}. \tag{49}$$

Translated into terms of G, (6) and (7) become

$$\begin{aligned} &(50) \qquad G''(x) = o([G'(x)]^{\frac{3}{2}}) \\ &(51) \qquad G'''(x) = o([G'(x)]^{2}). \end{aligned} \qquad (x \to \infty)$$

In view of (50) and (51) it follows from (48) and (49) that

$$\frac{\varphi''(x)}{\varphi'(x)} \to 0 \qquad (x \to \infty) \tag{52}$$

and

$$\frac{\varphi'''(x)}{\varphi'(x)} \to 0 \qquad (x \to \infty). \tag{53}$$

Now if $| z - v | <$ const., then

$$\frac{\varphi'(z)}{\varphi'(v)} \to 1 \qquad (z \to \infty) \tag{54}$$

since

$$\log \frac{\varphi'(z)}{\varphi'(v)} = \int_v^z \frac{\varphi''(x)}{\varphi'(x)}\, dx \to 0 \qquad (z \to \infty). \tag{55}$$

The relations (52), (53) and (54) yield the result (A).

For the proof[3] of (B) let us differentiate $A(u, \xi)$ partially with respect to u,

$$\begin{aligned} \frac{\partial A}{\partial u} = -\frac{\partial A}{\partial \psi(w)} = \frac{G(w) - \xi}{\psi'(w)} &= \frac{\varphi(\psi(w)) - \varphi(u + \psi(w))}{\varphi'(\psi(w))} \\ &= -u\, \frac{\varphi'(\theta u + \psi(w))}{\varphi'(\psi(w))}. \end{aligned} \tag{56}$$

[3] This proof has been revised in accordance with a suggestion of J. J. Gergen.

Now by (54),

$$\frac{\partial A}{\partial u} \to -u \qquad (\xi \to \infty) \tag{57}$$

uniformly for $|u| \leqq 2$. Next let ξ_0 be determined so that

$$-u - \tfrac{1}{2} \leqq \frac{\partial A}{\partial u} \leqq -u + \tfrac{1}{2}; \qquad A(u, \xi) \leqq -\tfrac{1}{2}u^2 + 1, \tag{58}$$

for $|u| \leqq 2$, $\xi_0 \leqq \xi$ and let w_0 be so determined that

$$|G''(w)| \leqq [G'(w)]^{\frac{3}{2}} \tag{59}$$

for $\psi(w) \geqq \psi(w_0)$. We shall now show that (B) holds with these values ξ_0, w_0. In fact, let $\xi_0 \leqq \xi$ and let $\bar{u}$ be any value of u such that

$$\bar{u} \leqq \varphi^{-1}(\xi) - \psi(w_0) \tag{60}$$

and

$$\left| \frac{\partial A}{\partial u}\right]_{u=\bar{u}} \Bigg| \leqq 1. \tag{61}$$

Then, since

$$\frac{\partial^2 A}{\partial u^2} = -1 + \tfrac{1}{2} \frac{G''(w)}{[G'(w)]^{\frac{3}{2}}} \frac{\partial A}{\partial u}, \tag{62}$$

it follows that

$$-\tfrac{3}{2} \leqq \left.\frac{\partial^2 A}{\partial u^2}\right]_{u=\bar{u}} \leqq -\tfrac{1}{2}. \tag{63}$$

In view of (56) it follows that

$$\frac{\partial A}{\partial u} \begin{array}{ll} > 0 & \text{if } u < 0, \\ < 0 & \text{if } u > 0, \end{array} \tag{64}$$

and hence

$$\left| \frac{\partial A}{\partial u} \right| < 1 \qquad \text{for } u \text{ on the open interval } (0, \bar{u}). \tag{65}$$

Since

$$\left.\frac{\partial A}{\partial u}\right]_{u=\frac{3}{2}} \leqq -1, \qquad \left.\frac{\partial A}{\partial u}\right]_{u=-\frac{3}{2}} \geqq 1, \tag{66}$$

it follows that

$$|\bar{u}| \leqq \tfrac{3}{2}. \tag{67}$$

Accordingly,

$$\frac{\partial A}{\partial u} < -1 \qquad \text{for } \xi_0 \leqq \xi, \quad 2 \leqq u \leqq \varphi^{-1}(\xi) - \psi(w_0); \tag{68}$$

$$\frac{\partial A}{\partial u} > 1 \qquad \text{for } \xi_0 \leqq \xi, \quad u \leqq -2, \; u \leqq \varphi^{-1}(\xi) - \psi(w_0). \tag{69}$$

(B) then follows.

Setting

$$p(\xi) = \varphi^{-1}(\xi) - \psi(w_0), \tag{70}$$

we see that (A) and (B) imply not only that for any positive λ

$$\sum_{\nu=-\infty}^{[p(\xi)]} \max_{\nu-\lambda \leqq u \leqq \nu} [e^{w\xi - F(\xi) - wG(w) + F(G(w))}] < M_\lambda < \infty \tag{71}$$

for all sufficiently large ξ, but also that for any positive λ

$$\lim_{\xi\to\infty} \sum_{\nu=-\infty}^{[p(\xi)]} \max_{\nu-\lambda \leqq u \leqq \nu} | e^{-\frac{1}{2}u^2} - e^{w\xi - F(\xi) - wG(w) + F(G(w))} | = 0, \tag{72}$$

for the ends of each function run down, as we have seen, faster than $e^{-\frac{1}{2}|u|}$ and any finite portion of

$$\exp \{w\xi - F(\xi) - wG(w) + F(G(w))\} \tag{73}$$

approaches $e^{-\frac{1}{2}u^2}$ as $\xi \to \infty$.

Let us consider

$$\sum_{w_0}^{\infty} a_n e^{n\xi - F(\xi)} = \sum_{w_0}^{\infty} a_n e^{nG(n) - F(G(n))} e^{n\xi - F(\xi) - nG(n) + F(G(n))}. \tag{74}$$

Since this approaches s as $\xi \to \infty$, it follows that the sum is $\leqq s + 1$ for $\xi \geqq \xi_2$ for some ξ_2. But $\exp \{n\xi - F(\xi) - nG(n) + F(G(n))\}$ is greater than some constant $h_\lambda > 0$ for $0 \leqq u \leqq \lambda$, $\xi \geqq \xi_3$, for some $\xi_3 \geqq \xi_2$, where $\xi = \varphi(u + \psi(n))$. Hence

$$\sum_{0 \leqq u \leqq \lambda} a_n e^{nG(n) - F(G(n))} \leqq \frac{s+1}{h_\lambda}, \qquad \xi \geqq \xi_3. \tag{75}$$

Now $0 \leqq u \leqq \lambda$ is equivalent to

$$\varphi^{-1}(\xi) - \lambda \leqq \psi(n) \leqq \varphi^{-1}(\xi) \tag{76}$$

and hence

$$\sum_{\nu \leqq \psi(n) \leqq \nu+\lambda} a_n e^{nG(n) - F(G(n))} < M'_\lambda \qquad (\nu = 0, 1, 2, \cdots). \tag{77}$$

From these facts we prove immediately that

$$\sum_{n=w_0}^{\infty} a_n e^{nG(n) - F(G(n)) - \frac{1}{2}u^2} \to s \qquad (\xi \to \infty); \tag{78}$$

indeed

$$\begin{aligned} &\left| \sum_{w_0}^{\infty} a_n e^{nG(n) - F(G(n))} \{e^{-\frac{1}{2}u^2} - e^{n\xi - F(\xi) - nG(n) + F(G(n))}\} \right| \\ &\leqq \sum_{\nu=-\infty}^{[p(\xi)]} \left\{ \left[\max_{\nu-1 \leqq u \leqq \nu} | e^{-\frac{1}{2}u^2} - e^{n\xi - F(\xi) - nG(n) + F(G(n))} | \right] \right. \\ &\qquad \left. \cdot \sum_{\nu-1 \leqq u \leqq \nu} a_n e^{nG(n) - F(G(n))} \right\} \to 0 \qquad (\xi \to \infty), \end{aligned} \tag{79}$$

and this proves (78). If we define

$$(80) \qquad L(y) = \begin{cases} (2\pi)^{\frac{1}{2}} \displaystyle\sum_{\psi(w_0) \leqq \psi(n) < y} a_n e^{nG(n) - F(G(n))} & \text{for } y > \psi(w_0), \\ 0 & \text{for } y \leqq \psi(w_0), \end{cases}$$

then (78) is equivalent to

$$(81) \qquad \int_{-\infty}^{\infty} e^{-\frac{1}{2}(x-y)^2} dL(y) \to s \int_{-\infty}^{\infty} e^{-\frac{1}{2}u^2} du \qquad (x \to \infty).$$

Now let us apply the General Tauberian Theorem of Wiener (loc. cit., pp. 27, 28) to (81). The kernel $e^{-\frac{1}{2}u^2}$ has the non-vanishing Fourier transform $e^{-\frac{1}{2}x^2}$; the function $L(y)$ is a monotone function such that

$$(82) \qquad \int_{\nu}^{\nu+\lambda} dL(y) = (2\pi)^{\frac{1}{2}} \sum_{\nu \leqq \psi(n) \leqq \nu+\lambda} a_n e^{nG(n) - F(G(n))} < (2\pi)^{\frac{1}{2}} M'_\lambda$$

and hence if

$$(83) \qquad K_2(u) = \begin{cases} 1 & \text{for } 0 \leqq -u \leqq \lambda, \\ 0 & \text{otherwise,} \end{cases}$$

then

$$(84) \qquad \int_{-\infty}^{\infty} K_2(x-y) dL(y) \to s \int_{-\infty}^{\infty} K_2(u) du = \lambda s.$$

That is,

$$(85) \qquad \int_{x}^{x+\lambda} dL(y) \to \lambda s \qquad (x \to \infty),$$

or

$$(86) \qquad \frac{(2\pi)^{\frac{1}{2}}}{\lambda} \sum_{x \leqq \psi(n) \leqq x+\lambda} a_n e^{nG(n) - F(G(n))} \to s \qquad (x \to \infty).$$

This concludes the proof of Theorem 1.

3. For the proof of Theorem 2 we note that the class Σ of functions of the form (83) is such that there is no x for which for every function $K_2(u)$ of Σ

$$(87) \qquad \frac{1}{(2\pi)^{\frac{1}{2}}} \int_{-\infty}^{\infty} K_2(u) e^{iux} du = 0.$$

Thus (84) implies[4] (81) and this proves Theorem 2. This argument shows the validity of Theorem 2 even if (8) holds only for two values of λ whose ratio is irrational.

[4] See Wiener, loc. cit., p. 26.

4. For the proof of Theorem 3 let (4), (5) and (10) hold. Then for some positive ϵ

$$G'(F'(x)) = \frac{1}{F''(x)} > \frac{1}{[F'(x)]^{1+\epsilon}} \qquad x > x_\epsilon, \tag{88}$$

and hence for every positive ϵ

$$x^{-1-\epsilon} < G'(x); \tag{89}$$

$$x^{-\frac{1}{2}-\epsilon} < \psi'(x), \tag{90}$$

for sufficiently large values of x. From (90) it follows that for every positive ϵ,

$$x^{\frac{1}{2}-\epsilon} < \psi(x) \tag{91}$$

for sufficiently large values of x. This proves part (i) of Theorem 3.

For part (ii) let (4), (5) and (13) hold. In view of (5) and (13) it follows that $G(x)$ is $O(x^\epsilon)$ for every $\epsilon > 0$. Applying the Schwarz inequality to (38), we have

$$\begin{aligned}\psi(x) &\leqq \left[\int_a^x dx \int_a^x G'(x)dx\right]^{\frac{1}{2}} \\ &= O(x^{\frac{1}{2}}[G(x)]^{\frac{1}{2}}) \\ &= O(x^{\frac{1}{2}+\epsilon}) \qquad \epsilon > 0,\end{aligned} \tag{92}$$

which yields (14).

5. Let (15) be an integral function. The condition (16) gives

$$\frac{1}{2\pi}\int_0^{2\pi} |f(re^{i\theta})|^2 d\theta = \sum_0^\infty |b_n|^2 r^{2n} \sim se^{F(2\log r)} \qquad (r \to \infty). \tag{93}$$

If we set

$$a_n = |b_n|^2, \qquad \xi = \log r^2, \tag{94}$$

(93) becomes

$$\sum_0^\infty a_n e^{n\xi - F(\xi)} \to s \qquad (\xi \to \infty). \tag{95}$$

Applying Theorem 1 to this case, we see that the conclusions of Theorem 4 follow. Theorem 5 is an immediate consequence of Theorem 2. In order to obtain the result on gaps stated in Theorem 6, let us assume that (11), (12) and (14) hold for sufficiently large values of n and let us consider the large values of n for which

$$x \leqq \psi(n) \leqq x + \lambda. \tag{96}$$

Let $g(x)$ be the inverse function to $\psi(x)$. Then the values of n for which (96) holds are those for which

$$g(x) \leqq n \leqq g(x + \lambda). \tag{97}$$

But

$$\begin{aligned} g(x+\lambda) &= g(x) + \lambda g'(x+\theta\lambda) \qquad\qquad 0 \leqq \theta \leqq 1 \\ &= g(x) + \lambda \frac{1}{\psi'(g(x+\theta\lambda))} \\ &= g(x) + \lambda O([g(x+\theta\lambda)]^{\frac{1}{2}+\epsilon}) \\ &= g(x) + \lambda O([g(x)]^{\frac{1}{2}+\epsilon}). \end{aligned} \tag{98}$$

Thus the values of n for which (96) holds are those for which

$$g(x) \leqq n \leqq g(x) + \lambda O([g(x)]^{\frac{1}{2}+\epsilon}) \tag{99}$$

and hence if $s \neq 0$, the function (15) can have only a finite number of gaps of magnitude $(\nu, \nu + \nu^{\frac{1}{2}+\epsilon})$.

6. Let $a_n \geqq 0$ $(n = 1, 2, \cdots)$, and let

$$c_n = \sum_1^n a_m a_{n-m} \qquad (n = 1, 2, \cdots). \tag{100}$$

If (8) holds for every positive value of λ, in view of Theorem 2 it follows that (2) holds and hence

$$\left[\sum_1^\infty a_n e^{n\xi}\right]^2 e^{-2F(\xi)} \to s^2 \qquad (\xi \to \infty); \tag{101}$$

$$\sum_1^\infty c_n e^{n\xi - 2F(\xi)} \to s^2 \qquad (\xi \to \infty). \tag{102}$$

In view of Theorem 1, (102) implies that

$$\lim_{x\to\infty} \frac{(2\pi)^{\frac{1}{2}}}{\lambda} \sum_{x \leqq \psi(n) \leqq x+\lambda} c_n e^{nG_1(n) - F_1(G_1(n))} = s^2, \tag{103}$$

where

$$\begin{aligned} F_1(x) &= 2F(x); & G_1(x) &= G(\tfrac{1}{2}x); \\ G_1'(x) &= \tfrac{1}{2}G'(\tfrac{1}{2}x); & \psi_1'(x) &= \frac{1}{\sqrt{2}}\psi'(\tfrac{1}{2}x); \end{aligned} \tag{104}$$

$$\psi_1(x) = \sqrt{2}\,\psi(\tfrac{1}{2}x) + \text{const.}$$

Thus (103) is equivalent to (20). Similarly, if (20) holds for every positive value of λ, (102) holds and hence (2) holds. This shows the complete equivalence of the two statements contained in Theorem 7.

MASSACHUSETTS INSTITUTE OF TECHNOLOGY.

SUMMABILITY OF CONJUGATE DERIVED SERIES

By W. C. Randels

We consider a function $f(x)$ with period 2π and integrable over $(-\pi, \pi)$. The Fourier series of such a function is

$$\tfrac{1}{2}a_0 + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx),$$

where

$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x) \cos nx \, dx, \qquad b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x) \sin nx \, dx.$$

The conjugate derived series of $f(x)$ is

$$\sum_{n=1}^{\infty} n(a_n \cos nx + b_n \sin nx).$$

We define

$$\varphi(t) = f(x + t) + f(x - t) - 2f(x)$$

and

$$\Phi(t) = \frac{-1}{4\pi}\int_{t}^{\pi} \varphi(y) \csc^2 \tfrac{1}{2}y \, dy.$$

Throughout this paper we shall suppose that $\varphi(t)/t \subset L$ on $(-\pi, \pi)$. This implies that $\Phi(t) \subset L$, since

$$\int_0^{\pi} | \Phi(t) | \, dt \leqq \frac{1}{4\pi}\int_0^{\pi} dt \int_t^{\pi} | \varphi(y) \csc^2 \tfrac{1}{2}y | \, dy = \frac{1}{4\pi}\int_0^{\pi} dy \, | \varphi(y) \csc^2 \tfrac{1}{2}y | \int_0^{y} dt$$

$$= O\left(\int_0^{\pi} | \varphi(y) | \frac{dy}{y}\right).$$

If $\alpha > 1$ and

$$\frac{\alpha - 1}{t^{\alpha-1}}\int_0^{t} \Phi(y)(t - y)^{\alpha-2} \, dy \to S \quad \text{as} \quad t \to +0,$$

we say that

$$\text{conj. der. lim } \varphi(t) = S(R', \alpha).$$

If $0 < \alpha \leqq 1$, and for some $\beta > 1$

$$\frac{\beta - 1}{t^{\beta-1}}\int_0^{t} \Phi(y)(t - y)^{\beta-2} \, dy \to S \quad \text{as} \quad t \to +0,$$

Received June 5, 1935; in revised form, February 7, 1937. The author is indebted to the referee for suggestions, which materially simplified the proof.

we say that

$$\text{conj. der. lim } \varphi(t) = S(R', \alpha)$$

if

$$(1) \qquad \frac{1}{t^{\alpha}} \int_0^t \varphi(y) y^{-1} (t - y)^{\alpha - 1} \, dy \to 0 \quad \text{as} \quad t \to +0.$$

The series

$$\sum_{n=0}^{\infty} a_n$$

is said to be (C, α) summable to the sum S if

$$\frac{1}{A_m^{\alpha}} \sum_{n=0}^{m} A_{m-n}^{\alpha} a_n \to S \quad \text{as} \quad m \to \infty,$$

where

$$A_n^{\alpha} = \frac{\Gamma(n + \alpha + 1)}{\Gamma(\alpha + 1)\Gamma(n + 1)}.$$

The theorems which we propose to prove are:

THEOREM I.[1] *If*

$$(2) \qquad \text{conj. der. lim } \varphi(t) = S(R', \alpha) \qquad (\alpha > 0),$$

then the conjugate derived series of $f(x)$ is (C, β) summable to the sum S for $\beta > \alpha + 1$.

THEOREM II. *If the conjugate derived series of $f(x)$ is (C, α), $\alpha \geqq 1$ to the sum S, then for $\beta > \alpha$*

$$\text{conj. der. lim } \varphi(t) = S(R', \beta).$$

These theorems are analogous to those proved for Fourier series by Paley[2] and Bosanquet,[3] for the conjugate series by Paley[4] and for the derived series by Takahashi.[5] These proofs used Riesz summability instead of Cesàro, but the two methods are known to be equivalent.

We may assume without loss of generality that the point under consideration is the point $x = 0$ and that the function is even. We also shall assume that

[1] Theorem I for $\alpha = 0$ has been proved by Gupta in a paper of a similar title, Proc. Acad. Sci. Allahabad, vol. 1 (1936), pp. 7–17. A generalization of this has been given by Moursund, Amer. Jour. Math., vol. 57 (1935), pp. 854–860. It might be mentioned that Bosanquet has considered the problem of the boundedness of the (C, α) transforms of the sequence $n(a_n \cos nx + b_n \sin nx)$ by similar methods, Trans. Amer. Math. Soc., vol. 39 (1936), pp. 189–205.

[2] R. E. A. C. Paley, Proc. Cam. Phil. Soc., vol. 26 (1930), pp. 173–203.

[3] L. S. Bosanquet, Proc. Lon. Math. Soc., vol. 31 (1930), pp. 144–164.

[4] L.c.

[5] T. Takahashi, Tohôku Math. Jour., vol. 38 (1933), pp. 265–278.

$a_t = 0$ and $f(0) = 0$. This is permissible, for otherwise we would deal with the function

$$f_0(t) = f(t) - a_0 - [f(0) + a_0] \cos t.$$

Then at $x = 0$, $\varphi(t) = 2f(t)$. We have

$$\Phi(t) + \frac{1}{4\pi}\int_t^\pi \varphi(y) \cot^2 \tfrac{1}{2}y\, dy = -\frac{1}{4\pi}\int_t^\pi \varphi(y)dy \to -\tfrac{1}{4}a_0 = 0 \quad \text{as} \quad t \to +0.$$

Hence, if $\alpha > 1$,

$$\frac{\alpha - 1}{t^{\alpha-1}} \int_0^t \Phi(y)(t - y)^{\alpha-2}\, dy \to S \quad \text{as} \quad t \to +0$$

implies that

$$\frac{\alpha - 1}{t^{\alpha-1}} \int_0^t \left\{\int_y^\pi \varphi(x) \cot^2 \tfrac{1}{2}x\, dx\right\}(t - y)^{\alpha-2}\, dy \to S \quad \text{as} \quad t \to +0.$$

Moreover it is clear that for $\alpha > 0$,

$$\frac{1}{t^\alpha}\int_0^t \frac{\varphi(y)}{y}(t - y)^{\alpha-1}\, dy \to 0 \quad \text{as} \quad t \to +0$$

implies that

$$\frac{1}{t^\alpha}\int_0^t \varphi(y) \cot \tfrac{1}{2}y\, (t - y)^{\alpha-1}\, dy \to 0 \quad \text{as} \quad t \to +0.$$

We set $g(t) = \varphi(t) \cot \frac{1}{2}t$ and, since $\varphi(t)/t \subset L$, $g(t) \subset L$. If $\psi(t) = g(x + t) - g(x - t)$, then in the notation of Paley (2) implies that

$$\text{(3)} \qquad \text{conj. lim } \psi(t) = 4S(R', \alpha).$$

We can easily see that conversely (2) implies (1). The Fourier coefficients of $g(t)$ are $a_n = 0$ and

$$\begin{aligned} b_n &= \frac{2}{\pi}\int_0^\pi \varphi(t) \cot \tfrac{1}{2}t \sin nt\, dt \\ &= \frac{2}{\pi}\int_0^\pi \varphi(t)\, \frac{\sin (n + \frac{1}{2})t}{\sin \frac{1}{2}t}\, dt - \frac{2}{\pi}\int_0^\pi \varphi(t) \cos nt\, dt. \end{aligned}$$

We are now in a position to prove our theorems. First by a theorem of Paley[6] (3) implies that for $\gamma > \alpha > 0$,

$$\frac{-1}{A_m^\gamma}\sum_{n=0}^{m} A_{m-n}^\gamma b_n \to 4S \quad \text{as} \quad m \to \infty.$$

But

$$\sum_{n=1}^{\infty} \frac{2}{\pi}\int_0^\pi \varphi(t) \cos nt\, dt = \varphi(0) = 0,$$

[6] L.c., Theorem II.

so that (3) implies that, for $\gamma > \alpha > 0$,

$$\frac{-1}{A_m^\gamma}\sum_{n=0}^{m} A_{m-n}^{\gamma} S_n = \frac{1}{A_m^\gamma}\sum_{n=0}^{m} A_{m-n}^{\gamma+1} a_n \to S \quad \text{as} \quad n \to \infty.$$

By a theorem of Andersen[7] this in turn implies that

$$\frac{1}{A_m^\beta}\sum_{n=0}^{m} A_{m-n}^{\beta} n a_n \to S \quad \text{as} \quad n \to \infty \qquad (\beta = \gamma + 1 > \alpha + 1).$$

This completes the proof of Theorem I.

Now let us suppose that

$$\frac{1}{A_m^\alpha}\sum_{n=0}^{m} A_{m-n}^{\alpha} n a_n \to S = 0 \quad \text{as} \quad n \to \infty \qquad (\alpha \geqq 1).$$

By the theorem of Andersen[8] used before, this implies that the series

$$\sum_{n=1}^{\infty} S_n$$

is $(C, \alpha - 1)$ summable to 0. By a theorem of Paley[9] this implies that if $\alpha \geqq 1$,

$$\text{conj. lim } \psi(t) = 0(R', \beta) \qquad (\beta > \alpha),$$

and as we have seen, this implies that

$$\text{conj. der. lim } \varphi(t) = 0(R', \beta) \qquad (\beta > \alpha),$$

and completes the proof of Theorem II.

It is necessary to use the condition $\alpha \geqq 1$ in Theorem II in order to apply the theorem of Paley. As a matter of fact, Paley's results hold for one lower index. Compare Bosanquet's result for the analogous problem in Fourier series.[10] I have a direct proof of Theorem II for $\alpha > 0$.

It might be mentioned that Theorem I is true for $\alpha = 0$ if

$$\text{conj. der. lim } \varphi(t) = S(R', 0)$$

is defined by replacing condition (1) by

$$\frac{\varphi(t)}{t} \to 0 \quad \text{as} \quad t \to +0.$$

NORTHWESTERN UNIVERSITY.

[7] A. F. Andersen, Proc. Lon. Math. Soc., vol. 27 (1927), pp. 39–71; see Theorem 3.

[8] L.c., Theorem 3.

[9] L.c., Theorem IV.

[10] L.c., Theorem 4.

ORTHOGONAL POLYNOMIALS ON A PLANE CURVE

By Dunham Jackson

1. **Introduction.** Polynomials in two real variables x and y orthogonal with respect to integration along a curve in the (x, y)-plane can be constructed by the usual process for building up a system of orthogonal functions. If the curve is not algebraic, they have formal properties closely corresponding to those of polynomials orthogonal over a two-dimensional region.[1] For an algebraic curve the relations are different in important respects, reverting in some degree toward those which are familiar in the case of orthogonal functions of a single variable. This is to be pointed out in detail below, under hypotheses which, though not of the utmost generality, are still sufficiently illustrative.

2. **Orthogonal polynomials on a non-algebraic curve.** Let $\varphi(t)$, $\psi(t)$ be continuous functions of t, of period A. If they are not both constant, the equations $x = \varphi(t)$, $y = \psi(t)$ may be regarded as defining a closed curve C (not necessarily of simple character). A relation of linear dependence connecting any finite number of the functions $[\varphi(t)]^h$, $[\psi(t)]^k$, $h = 0, 1, 2, \cdots, k = 0, 1, 2, \cdots$, would mean that a polynomial in x and y vanishes identically on the curve, and so that the curve is the locus or a part of the locus of an algebraic equation. *Let it be assumed for the present that no such relation of linear dependence exists.*

Let $\rho(t)$ be a non-negative integrable function of period A, which, if not everywhere positive, is at any rate such that its product with any polynomial in $\varphi(t)$ and $\psi(t)$ (having a non-vanishing coefficient) is different from zero for a set of values of t of positive measure in a period. This condition will be satisfied, for example, if there is an interval in which $\rho(t)$ is almost everywhere different from zero and for which the corresponding points (x, y) do not belong to an algebraic locus.

Under the hypotheses that have been formulated any finite number of the quantities $\rho^{\frac{1}{2}}, \rho^{\frac{1}{2}}x, \rho^{\frac{1}{2}}y, \rho^{\frac{1}{2}}x^2, \rho^{\frac{1}{2}}xy, \rho^{\frac{1}{2}}y^2, \cdots$, regarded as functions of t for $0 \leqq t \leqq A$, are linearly independent. It is possible by "Schmidt's process of orthogonalization" to construct from them a sequence of functions which are orthogonal and normalized over the interval. When the functions are taken in the order indicated, the members of the orthogonal set are of the form $[\rho(t)]^{\frac{1}{2}}q_{nm}(x, y)$, $n = 0, 1, 2, \cdots, m = 0, 1, \cdots, n$, where $x = \varphi(t)$, $y = \psi(t)$, and q_{nm} is a polynomial of degree n in x and y together, while m is the exponent

Received November 16, 1936; presented to the American Mathematical Society, September 3, 1936.

[1] Cf. D. Jackson, *Formal properties of orthogonal polynomials in two variables*, this Journal, vol. 2 (1936), pp. 423–434; referred to hereafter as paper A.

of the highest power of y occurring in a term of the n-th degree.[2] The q's will be said to constitute a set of normalized orthogonal polynomials on the curve C, with respect to t as parameter and $\rho(t)$ as weight function, satisfying the conditions that

$$\int_C \rho(t) q_{kl}(x, y) q_{nm}(x, y)\, dt = 0, \qquad |n - k| + |m - l| \neq 0,$$

$$\int_C \rho(t) [q_{nm}(x, y)]^2\, dt = 1.$$

Certain properties of these polynomials can be stated briefly, the proofs being so similar to those for the case of polynomials orthogonal over a region[3] that it is unnecessary to give them in detail; it is to be borne in mind that under the present hypotheses, if the product of ρ by a polynomial in x and y vanishes almost everywhere as a function of t on the interval $(0, A)$, the coefficients of the polynomial must all be zero. Any polynomial of the n-th degree which is orthogonal (in the sense under consideration) to every polynomial of lower degree must be a linear combination of $q_{n0}, \cdots, q_{nn}$. Any $n + 1$ polynomials $p_{n0}, \cdots, p_{nn}$ which are expressible in terms of $q_{n0}, \cdots, q_{nn}$ by an orthogonal linear transformation are orthogonal to each other, normalized, and orthogonal to every polynomial of lower degree. And any $n + 1$ polynomials of the n-th degree which are normalized, orthogonal to each other, and orthogonal to every polynomial of lower degree are related to any other such set, and in particular to $q_{n0}, \cdots, q_{nn}$, by an orthogonal transformation.

Suppose there is a number μ between 0 and A such that

$$\begin{aligned} \varphi(t + \mu) &= \alpha\varphi(t) + \beta\psi(t), \\ \psi(t + \mu) &= \gamma\varphi(t) + \delta\psi(t), \\ \rho(t + \mu) &= \rho(t) \end{aligned} \tag{1}$$

for all values of t, where $\alpha, \beta, \gamma, \delta$ are constants. The determinant $\alpha\delta - \beta\gamma$ is certainly different from zero; otherwise there would be a linear relation connecting $\varphi(t + \mu)$ and $\psi(t + \mu)$, in contradiction with the hypothesis that the curve is not algebraic. If (x, y) is a point of the curve, the point (x', y') with coördinates given by

$$x' = \alpha x + \beta y, \qquad y' = \gamma x + \delta y \tag{2}$$

is also a point of the curve, and vice versa; the transformation (2) carries the curve into itself.

[2] In a corresponding passage in the paper A, p. 424, "the degree with respect to y" should be amended to read "the degree of the leading homogeneous aggregate of terms with respect to y".

[3] See the paper A, pp. 424–425.

If $P(x, y)$ is any polynomial in x and y,

$$\int_0^A \rho(t)P[\alpha\varphi(t) + \beta\psi(t), \gamma\varphi(t) + \delta\psi(t)]\,dt$$

$$= \int_{t=0}^{t=A} \rho(t + \mu)P[\varphi(t + \mu), \psi(t + \mu)]\,d(t + \mu)$$

$$= \int_\mu^{A+\mu} \rho(t)P[\varphi(t), \psi(t)]\,dt$$

$$= \int_0^A \rho(t)P[\varphi(t), \psi(t)]\,dt,$$

the last equality resulting from the periodicity of the functions involved. In summary, if $P(x', y')$ as a polynomial in x and y is denoted by $\Pi(x, y)$,

$$\int_C \rho(t)\Pi(x, y)\,dt = \int_C \rho(t)P(x', y')\,dt = \int_C \rho(t)P(x, y)\,dt. \tag{3}$$

For any polynomial $p(x, y)$ which is normalized on C, let $p(x', y') \equiv \pi(x, y)$. Then the last relation, with $P(x, y) \equiv [p(x, y)]^2$, says that $\pi(x, y)$ also is normalized. If $p(x, y)$ and $s(x, y)$ are two polynomials which are orthogonal to each other, and if $p(x', y') \equiv \pi(x, y)$, $s(x', y') \equiv \sigma(x, y)$, the relation (3) with $P \equiv ps$ means that $\pi(x, y)$ and $\sigma(x, y)$ are orthogonal. If $p(x, y)$ is of the n-th degree, and the condition of orthogonality is satisfied for every polynomial $s(x, y)$ of lower degree, then $\sigma(x, y)$ also is an arbitrary polynomial of degree lower than the n-th, and $\pi(x, y)$ is orthogonal to every such polynomial. Applied in particular to any set of $n + 1$ polynomials $p_{n0}, \cdots, p_{nn}$ of the n-th degree which are normalized, orthogonal to each other, and orthogonal to every polynomial of lower degree, this reasoning shows that $p_{n0}(x', y'), \cdots, p_{nn}(x', y')$ are expressible in terms of $p_{n0}(x, y), \cdots, p_{nn}(x, y)$ by an orthogonal linear transformation.

The work of the last three paragraphs applies also, with obvious modifications in detail, if the equations (1) are replaced by

$$\begin{aligned} \varphi(\mu - t) &= \alpha\varphi(t) + \beta\psi(t), \\ \psi(\mu - t) &= \gamma\varphi(t) + \delta\psi(t), \\ \rho(\mu - t) &= \rho(t). \end{aligned} \tag{4}$$

Furthermore, the entire discussion remains valid if the functions φ, ψ, ρ are defined merely for $0 \leqq t \leqq A$, the hypothesis of periodicity being dropped, and the closed curve C replaced by a non-algebraic arc, except that in place of (1) and (4) only the equations

$$\begin{aligned} \varphi(A - t) &= \alpha\varphi(t) + \beta\psi(t), \\ \psi(A - t) &= \gamma\varphi(t) + \delta\psi(t), \\ \rho(A - t) &= \rho(t) \end{aligned}$$

come into consideration. Other extensions readily suggest themselves, though it is not so clear what the most general possible formulation would be.

Under conditions of some generality the transformation (2) must itself be orthogonal. Let the curve C be rectifiable, and let it be represented by the equations

$$x = \varphi(s), \qquad y = \psi(s)$$

in terms of the arc-length as parameter. For simplicity, let it be made up of a finite number of arcs, on each of which $\varphi(s)$ and $\psi(s)$ have continuous derivatives. Then

$$[\varphi'(s)]^2 + [\psi'(s)]^2 = 1 \tag{5}$$

except for a finite number of points at most. Let φ and ψ satisfy identically the equations

$$\varphi(\mu + s) = \alpha\varphi(s) + \beta\psi(s),$$
$$\psi(\mu + s) = \gamma\varphi(s) + \delta\psi(s),$$

or the alternative equations with $\mu + s$ replaced by $\mu - s$. Then

$$\begin{aligned} 1 &= [\varphi'(\mu \pm s)]^2 + [\psi'(\mu \pm s)]^2 \\ &= (\alpha^2 + \gamma^2)[\varphi'(s)]^2 + 2(\alpha\beta + \gamma\delta)\varphi'(s)\psi'(s) + (\beta^2 + \delta^2)[\psi'(s)]^2, \end{aligned}$$

which with (5) means that

$$(\alpha^2 + \gamma^2 - 1)[\varphi'(s)]^2 + 2(\alpha\beta + \gamma\delta)\varphi'(s)\psi'(s) + (\beta^2 + \delta^2 - 1)[\psi'(s)]^2 = 0,$$

except at a finite number of points. Let the slope $dy/dx = \psi'(s)/\varphi'(s)$ be denoted by λ. If λ takes on three or more distinct finite values, the quadratic equation

$$(\alpha^2 + \gamma^2 - 1) + 2(\alpha\beta + \gamma\delta)\lambda + (\beta^2 + \delta^2 - 1)\lambda^2 = 0$$

has more than two distinct roots, and its coefficients must vanish; the equations

$$\alpha^2 + \gamma^2 = 1, \qquad \beta^2 + \delta^2 = 1, \qquad \alpha\beta + \gamma\delta = 0$$

are the conditions of orthogonality. The same conclusion holds if $\varphi'(s)/\psi'(s)$ takes on three distinct finite values, and so if the slope takes on the value ∞ and two finite values different from zero; and again if there are regular points of the curve for which respectively $\varphi'(s) = 0$, $\psi'(s) = 0$, and $\varphi'(s)\psi'(s) \neq 0$, i.e., if the slope takes on the values 0 and ∞ and a finite value different from 0.

These brief observations raise numerous questions which are left unanswered. It may be pointed out, however, that the hypothesis that the curve has at least three different directions is not entirely irrelevant. For the parallelogram with vertices at the points $(\pm 2, 0)$ and $(0, \pm 1)$ (an algebraic curve, to be sure) is carried over into itself by the non-orthogonal transformation $x' = -2y$, $y' = \frac{1}{2}x$, arising from the substitution of $s + 5^{\frac{1}{2}}$ for s.

The pertinent parts of the discussion in the paper A relating to particular

transformations (2) and the orthogonal transformations of $p_{n0}, \cdots, p_{nn}$ induced by them, specifically in the case of the transformations $x' = -x, y' = -y$ (see footnote 6 of paper A) and $x' = y, y' = x$, can be carried over to the present situation. Also the reasoning of the second section of that paper, which is concerned with a recursion formula and a Christoffel-Darboux identity for the orthogonal polynomials, is applicable here with the obvious formal adaptations.

3. **Definition and orthogonal transformation of orthogonal polynomials on an algebraic curve.** The foregoing conclusions have to be modified if the curve to which the discussion relates is algebraic. Let the weight function $\rho(t)$ for simplicity be taken as positive everywhere or almost everywhere on its range of definition, so that questions of linear dependence shall not be complicated by the possibility of vanishing of ρ. Let $\varphi(t)$ and $\psi(t)$ once more be continuous for $0 \leqq t \leqq A$, or continuous everywhere and of period A, and not both constant, but let it be supposed now that there is a polynomial in x and y which vanishes identically for $0 \leqq t \leqq A$ when x and y are replaced by $\varphi(t)$ and $\psi(t)$. Any multiple of such a polynomial will naturally have the same property. If a factor of a polynomial of this sort vanishes at only a finite number of points of the curve C, the quotient obtained by dividing out this factor will vanish at points arbitrarily near the finite number of points in question, and will vanish at those points by continuity and so at all points of the curve. There is therefore a polynomial vanishing identically on C, and composed of irreducible factors, each of which vanishes at infinitely many points of C.

If $\Omega_1(x, y)$ is a particular polynomial meeting these specifications, any irreducible polynomial vanishing at infinitely many points of C must vanish at infinitely many points simultaneously with one of the irreducible factors of $\Omega_1(x, y)$, and must be identical with that factor except for a constant multiplier.[4] There can be only a finite number of essentially distinct irreducible polynomials vanishing at infinitely many points of C; their product, determined except for a constant factor, may be characterized as the polynomial of lowest degree vanishing identically on C. Let this polynomial be denoted by $\Omega(x, y)$, and let its degree in the two variables together be N; the choice of the constant multiplier is immaterial. (The irreducible factors are essentially real; if an irreducible polynomial with complex coefficients vanishes at infinitely many real points, its conjugate, vanishing at the same real points, must be a constant multiple of it.) Any polynomial which vanishes identically on C is a multiple of $\Omega(x, y)$. The curve C is by no means necessarily the complete locus of the equation $\Omega(x, y) = 0$; it may be a triangle or a square, or otherwise composed of a number of algebraic arcs.

Let the *rank* of a monomial $x^h y^k$ for the moment denote the index of its position in the sequence $1, x, y, x^2, xy, y^2, \cdots$, so that if two terms are of different degrees in the two variables together, the one of higher degree has the higher

[4] See for example, M. Bôcher, *Introduction to Higher Algebra*, pp. 210–211.

rank, while if two terms are of the same degree the one which is of higher degree in y has the higher rank. If the terms of two polynomials are arranged according to rank in descending order, the leading term of the product of the two polynomials is the product of their leading terms. If the leading term of $\Omega(x, y)$ is $x^p y^q$, $p + q = N$, the leading term of any polynomial which vanishes identically on C must be divisible by $x^p y^q$. And if h and k are any two exponents such that $h \geqq p$, $k \geqq q$, multiplication of $\Omega(x, y)$ by $x^{h-p} y^{k-q}$ gives a polynomial vanishing identically on C, and having $x^h y^k$ for its leading term. In other words, among the monomials $x^n, x^{n-1}y, \cdots, y^n$ of the n-th degree, when $n \geqq N$, each member of the set $x^{n-q}y^q, x^{n-q-1}y^{q+1}, \cdots, x^p y^{n-p}$ is linearly expressible on the curve C in terms of monomials of lower rank than its own, while none of the remaining N monomials of the n-th degree, in which the first exponent is less than p or the second exponent less than q, is connected with terms of lower rank than its own by any relation of linear dependence on C. If all terms $x^h y^k$ in which simultaneously $h \geqq p$ and $k \geqq q$ are omitted from the sequence 1, $x, y, x^2, xy, y^2, \cdots$, Schmidt's process can be applied to the remaining terms, taken in the order indicated, and yields a set of normalized orthogonal polynomials comprising just N polynomials of the n-th degree for each value of $n \geqq N$ (and $n + 1$ polynomials of degree n for $n < N$).

The preceding argument as it stands is of course based essentially on a particular arbitrary arrangement of the monomials $x^h y^k$ in serial order. It remains to be seen to what extent it possesses more general significance.

In the orthogonal system that has been described let the polynomials of the n-th degree be denoted by $q_{n1}, q_{n2}, \cdots, q_{nN}$, for $n \geqq N$, where the first subscript indicates the degree in the two variables together, and for fixed n the second subscript increases with the exponent of the highest power of y occurring in a term of the n-th degree, but is not in general equal to that exponent. The modifications required for values of $n < N$ will be obvious. Let the leading term of a polynomial be characterized still in terms of the notion of rank already employed. Each of the monomials $x^n, x^{n-1}y, \cdots, y^n$, with an appropriate coefficient, is leading term either of one of the polynomials q_{nk} or of a polynomial which vanishes identically on C. If $p(x, y)$ is an arbitrary polynomial of the n-th degree orthogonal to every polynomial of lower degree, suitable constant multiples of the $n + 1$ polynomials just mentioned can be subtracted from $p(x, y)$ to remove successively all the terms of the n-th degree in order of decreasing rank, leaving a polynomial which, being orthogonal to every polynomial of degree lower than the n-th, must be orthogonal to itself and so identically zero on the curve. This means that $p(x, y)$ is equal on the curve to a linear combination of $q_{n1}, \cdots, q_{nN}$, or as a polynomial in x and y is the sum of such a linear combination and a polynomial which contains $\Omega(x, y)$ as a factor.

It follows that any set of N polynomials $p_{n1}, \cdots, p_{nN}$ of the n-th degree which are normalized, orthogonal to each other, and orthogonal to every polynomial of lower degree, can be expressed on C in terms of $q_{n1}, \cdots, q_{nN}$ by an orthogonal transformation. Clearly any set of N polynomials thus expressible

in terms of the q's will have the properties mentioned. And since for points on the curve the q's are given in terms of the p's by the inverse orthogonal transformation, any polynomial of the n-th degree orthogonal to every polynomial of lower degree can be linearly represented on the curve in terms of the p's, and any two sets of p's are related to each other on the curve by an orthogonal transformation. It is perhaps not superfluous to emphasize that all these relations, which for the polynomials in x and y have the character of congruences with respect to $\Omega(x, y)$ as modulus, are identities in the functions of t which are obtained on replacement of x and y by $\varphi(t)$ and $\psi(t)$. In terms of x and y the individual polynomials of the n-th degree in the orthogonal system are subject not merely to the indeterminacy involved in the admissibility of orthogonal transformation, but also to the addition of arbitrary polynomials of the n-th degree containing $\Omega(x, y)$ as a factor.

If the curve C is a straight line segment, $N = 1$, and the orthogonal system contains just one polynomial of each degree. With s as parameter and weight function unity the orthogonal polynomials are essentially the Legendre polynomials in s, and introduction of a weight function $\rho(s)$ gives rise to the corresponding set of orthogonal polynomials in a single variable.

If C is the unit circle, with the parametric representation

$$x = \cos t, \qquad y = \sin t, \qquad 0 \leqq t \leqq 2\pi,$$

and $\rho(t) = 1/\pi$, a polynomial in x and y is a trigonometric sum in t, the orthogonal polynomials of the n-th degree ($N = 2$), chosen with the exercise of a particular option as regards orthogonal transformation, reduce on the curve to $\cos nt$ and $\sin nt$ for $n \geqq 1$, and in terms of x and y they are any polynomial representations of $\cos nt$ and $\sin nt$ in terms of $\cos t$ and $\sin t$, e.g., a polynomial of the n-th degree in x and the product of y by a polynomial of degree $n - 1$ in x, or, alternatively, the real and pure imaginary parts of $(x + iy)^n$. A general orthogonal transformation on $\cos nt$ and $\sin nt$ has the effect merely of replacing them by $\cos (nt + k)$ and $\pm \sin (nt + k)$. Admission of a non-constant weight function $\rho(t)$ leads to more general orthogonal trigonometric sums of the type discussed in a recent paper by the writer.[5]

Substitution of the ellipse $x = a \cos t$, $y = b \sin t$ for the circle, with a given $\rho(t)$, does not change the orthogonal functions as to their dependence on t, but changes the coefficients of their representation in terms of x and y, by the substitution of x/a and y/b for x and y. Solution of the problem for the ellipse with s as parameter and constant weight function would correspond to the construction of orthogonal trigonometric sums in t (not trigonometric sums in s) with $\rho(t) = ds/dt$.

The rather obvious remarks of the last three paragraphs are inserted merely for the sake of showing more clearly how the theory under discussion may be

[5] D. Jackson, *Orthogonal trigonometric sums*, Annals of Mathematics, vol. 34 (1933), pp. 799–814.

regarded as a generalization of that of the most familiar orthogonal systems in one dimension.

If φ, ψ, ρ satisfy a set of relations (1) or (4), and if x', y' are defined by (2), the reasoning carried through for a non-algebraic curve can be used now to show that $p_{n1}(x', y'), \cdots, p_{nN}(x', y')$ are expressible *on the curve* in terms of $p_{n1}(x, y), \cdots, p_{nN}(x, y)$ by an orthogonal transformation, provided that C is not a straight line segment. In the excluded case the inference as to the non-vanishing of $\alpha\delta - \beta\gamma$ is inadmissible; there is perhaps no need of dwelling further on this case in the two-dimensional setting, the facts with regard to orthogonal polynomials in a single variable being well known. In particular, if (2) is the transformation $x' = -x$, $y' = -y$, $p_{ni}(x', y')$ is identically equal on the curve to $(-1)^n p_{ni}(x, y)$, for $i = 1, 2, \cdots, N$. But the transformation $x' = y$, $y' = x$ requires further consideration, which will not be undertaken here; for in the two-dimensional case,[6] or in the case of a non-algebraic curve, it is possible to say that the polynomial $q_{n0}(x, y)$, for example, is neither symmetric nor skew-symmetric, while on the circle $x^2 + y^2 = 1$, with constant weight function, the corresponding polynomial for $n = 2$ satisfies the congruence-identity

$$2x^2 - 1 \equiv -(2y^2 - 1).$$

4. **Recursion formula and Christoffel-Darboux identity on an algebraic curve.** By reason of the fact that the number of polynomials of the n-th degree in the orthogonal system does not increase indefinitely with n, the recursion formula and the Christoffel-Darboux identity have a somewhat closer resemblance to those found in the case of a single variable than when the domain of orthogonality is a non-algebraic curve or a two-dimensional region. It is to be kept in mind, however, that the various identities hold in general only on the curve.

Let $p_{ni}(x, y)$ be an arbitrary one of the polynomials of the n-th degree. The product $xp_{ni}(x, y)$, as a polynomial of degree $n + 1$, is expressible on the curve in the form

$$(6) \qquad xp_{ni}(x, y) = \textstyle\sum_j A_{nij}p_{n+1,j}(x, y) + \sum_j B_{nij}p_{nj}(x, y) + \sum_j C_{nij}p_{n-1,j}(x, y),$$

with

$$A_{nij} = \int_C \rho(t)xp_{ni}(x, y)p_{n+1,j}(x, y)\,dt,$$

$$B_{nij} = \int_C \rho(t)xp_{ni}(x, y)p_{nj}(x, y)\,dt = B_{nji},$$

$$C_{nij} = \int_C \rho(t)xp_{ni}(x, y)p_{n-1,j}(x, y)\,dt = A_{n-1,ji}.$$

Terms of lower degree are absent from the right-hand member because the corresponding coefficients vanish, by reason of the property of orthogonality.

[6] See paper A, p. 427.

In this section, in the absence of express indication to the contrary, the sign Σ always denotes summation over the designated index from 1 to N. The expression (6) is then appropriate in the first instance for values of $n \geqq N$; it can be made formally valid for $0 \leqq n \leqq N - 1$, however, by adoption of the convention, in the integral formulas for the coefficients as well as in the identity itself, that $p_{nj}(x, y) \equiv 0$ when $0 \leqq n + 1 < j \leqq N$. With this convention, which will be maintained henceforth, the identity holds trivially when $1 \leqq n + 1 < i \leqq N$ in the left-hand member, as well as significantly when $p_{ni}(x, y)$ is not identically zero. The recursion formula (6) resembles that for a single variable in that the number of its terms does not increase with n.

Multiplication of (6) by $p_{ni}(u, v)$, followed by summation over the index i from 1 to N, with replacement of the coefficient C_{nij} by its equal $A_{n-1,ji}$, gives

$$
\begin{aligned}
x \textstyle\sum_i p_{ni}(x, y)p_{ni}(u, v) &= \textstyle\sum_i \sum_j A_{nij} p_{n+1,j}(x, y)p_{ni}(u, v) \\
&+ \textstyle\sum_i \sum_j B_{nij} p_{nj}(x, y)p_{ni}(u, v) + \sum_i \sum_j A_{n-1,ji} p_{n-1,j}(x, y)p_{ni}(u, v).
\end{aligned}
\tag{7}
$$

On interchange of (x, y) with (u, v) this becomes

$$
\begin{aligned}
u \textstyle\sum_i p_{ni}(x, y)p_{ni}(u, v) &= \textstyle\sum_i \sum_j A_{nij} p_{n+1,j}(u, v)p_{ni}(x, y) \\
&+ \textstyle\sum_i \sum_j B_{nij} p_{nj}(u, v)p_{ni}(x, y) + \sum_i \sum_j A_{n-1,ji} p_{n-1,j}(u, v)p_{ni}(x, y).
\end{aligned}
\tag{8}
$$

In the second summation on the right-hand side of (8) let the index symbols i and j be interchanged, the coefficient B_{nij} being restored, however, by virtue of the fact that $B_{nij} = B_{nji}$; the sum in question then becomes identical with the corresponding sum in (7). Subtraction of (7) from (8), with interchange of i and j in the last summation of each identity, gives

$$
\begin{aligned}
&(u - x) \textstyle\sum_i p_{ni}(x, y)p_{ni}(u, v) \\
&\qquad = \textstyle\sum_i \sum_j A_{nij}[p_{n+1,j}(u, v)p_{ni}(x, y) - p_{n+1,j}(x, y)p_{ni}(u, v)] \\
&\qquad - \textstyle\sum_i \sum_j A_{n-1,ij}[p_{nj}(u, v)p_{n-1,i}(x, y) - p_{nj}(x, y)p_{n-1,i}(u, v)].
\end{aligned}
\tag{9}
$$

For the evaluation of

$$K_n(x, y, u, v) \equiv \sum_{k=0}^{n} \sum_{i=1}^{N} p_{ki}(x, y)p_{ki}(u, v)$$

let (9) be written with k in place of n for each value of k from 0 to n. The result of summation with respect to k, when due account is taken of the fact that $p_{-1,i} \equiv 0$ for all values of i, is

$$(u - x)K_n(x, y, u, v) = \textstyle\sum_i \sum_j A_{nij}[p_{n+1,j}(u, v)p_{ni}(x, y) - p_{n+1,j}(x, y)p_{ni}(u, v)].$$

In this relation, which has the character of a Christoffel-Darboux identity, the number of terms again remains constant as n increases. There is naturally a corresponding formula for $(v - y)K_n(x, y, u, v)$, differing from this only in the values of the coefficients A_{nij}.

The University of Minnesota.

THE CLASSES OF INTEGRAL SETS IN A QUATERNION ALGEBRA

By Claiborne G. Latimer

1. **Introduction.** Let $\mathfrak{A}$ be a rational generalized quaternion algebra with the fundamental number d.[1] A set of integral elements in $\mathfrak{A}$, or more briefly an integral set, is one with certain properties R, C, U, M as defined by Dickson.[2] Two integral sets are said to be equivalent, or of the same type, if there is a one-to-one correspondence between the elements of the sets which is preserved under addition and multiplication. All the sets equivalent to a given set will be said to form a class. Two integral sets, $\mathfrak{G}$ and $\mathfrak{G}_1$, belong to the same class if and only if there is a non-singular element α in $\mathfrak{A}$ such that $\mathfrak{G}_1 = \alpha\mathfrak{G}\alpha^{-1}$.[3]

By a result due to Artin,[4] the number H of classes of integral sets in $\mathfrak{A}$ is equal to the number of classes of equivalent right ideals in an arbitrarily chosen integral set $\mathfrak{G}$, of $\mathfrak{A}$, Artin's definition of equivalent ideals being broader than the usual definition.

The principal purpose of this paper is to show that there is a one-to-one correspondence between the classes of integral sets in $\mathfrak{A}$ and certain classes of ternary quadratic forms. These classes of forms are the non-negative classes or the improperly primitive non-negative classes in a certain genus G, according as d is even or odd. G is uniquely determined by d. If $d < 0$, by a known theorem there is a single class of forms in G and therefore $H = 1$.[5]

We shall also determine a relatively simple basis of an arbitrarily chosen integral set in $\mathfrak{A}$.

2. **A normal basis of an integral set.** If $\lambda_0, \cdots, \lambda_3$ form a basis of an integral set $\mathfrak{G}$, then $\lambda_i\lambda_j = \Sigma_k c_{ijk}\lambda_k$ $(i, j = 0, \cdots, 3)$. A set $\mathfrak{G}_1$ is equivalent to $\mathfrak{G}$ if and only if it has a basis $\xi_0, \cdots, \xi_3$ such that $\xi_i\xi_j = \Sigma_k c_{ijk}\xi_k$ $(i, j = 0, \cdots, 3)$. The following theorem is a consequence of certain results due to Brandt.[6]

Theorem 1. *Let $\mathfrak{A}$ be a generalized quaternion algebra with the fundamental*

Received November 16, 1936.

[1] For the definition of d, see Brandt, *Idealtheorie in Quaternionenalgebren*, Mathematische Annalen, vol. 99 (1929), p. 9.

[2] *Algebras and their Arithmetics*, pp. 141, 2. It may be shown that our definition of an integral set is equivalent to Brandt's definition of a *maximaler Integritätsbereich*, loc. cit., p. 11.

[3] Deuring, *Algebren*, p. 89.

[4] Abhandlungen aus dem Mathematischen Seminar der Hamburgischen Universität, vol. 5 (1927), p. 288, Theorem 20.

[5] In another paper, it was shown that if $d < 0$, then every one-sided ideal in an integral set is principal. (Transactions of the American Mathematical Society, vol. 40 (1936), p. 322.) From this and Artin's result, cited above, it again follows that if $d < 0$, then $H = 1$.

[6] Loc. cit., pp. 8–11.

number d, and let $\lambda_0 = 1, \cdots, \lambda_3$ be a basis of $\mathfrak{A}$. The λ's form a basis of an integral set if and only if $N(\Sigma x_i\lambda_i) = \Phi(x_0, \cdots, x_3) = \frac{1}{2}\Sigma g_{ij}x_ix_j$, where

(a) *the coefficients of Φ are integers, $g_{ij} = g_{ji}$, $g_{00} = 2$;*

(b) *the determinant $|g_{ij}| = d^2$;*

(c) *every third order minor in the matrix (g_{ij}) is divisible by d and every principal third order minor is divisible by $2d$.*

Two integral sets in $\mathfrak{A}$ are equivalent if and only if they have basal elements $\lambda_0, \cdots, \lambda_3$ and $\xi_0, \cdots, \xi_3$ respectively, such that $N(\Sigma x_i\lambda_i) \equiv N(\Sigma x_i\xi_i)$.

Let $\mathfrak{G}$ be an integral set in $\mathfrak{A}$ with the basis $\lambda_0 = 1, \cdots, \lambda_3$ and let $N(\Sigma x_i\lambda_i) = \frac{1}{2}\Sigma g_{ij}x_ix_j$ as in Theorem 1. Suppose every g_{0i} were even. Then d would be even and, since every $g_{ii} = 2g'_{ii}$ is even, by (c) of Theorem 1 every g_{ij} would be even. Then d would be divisible by 4, whereas d contains no square factor > 1.[7] Therefore one of the g_{0i} is odd. We may then assume, after an integral transformation of determinant unity, that $g_{01} = 1$, $g_{02} = g_{03} = 0$.

The trace, or double the scalar part, of $X = \Sigma x_i\lambda_i$ is $T(X) = 2x_0 + x_1$.[8] Then $T(\lambda_1) = 1$, $T(\lambda_2) = T(\lambda_3) = 0$. Such a basis of $\mathfrak{G}$ will be called a normal basis. For a normal basis, the matrix (g_{ij}) of 2Φ is in the form

$$(1) \qquad M = \begin{pmatrix} 2 & 1 & 0 & 0 \\ 1 & g_{11} & g_{12} & g_{13} \\ 0 & g_{21} & g_{22} & g_{23} \\ 0 & g_{31} & g_{32} & g_{33} \end{pmatrix}.$$

Since the determinant $|M| = d^2$, and every g_{ii} is even, it follows that $g_{23} \equiv d \pmod 2$.

3. The class of ternary forms corresponding to a class of integral sets. Let $\lambda_0 = 1, \cdots, \lambda_3$ be a normal basis of the integral set $\mathfrak{G}$, and let $\mathfrak{F}$ be the module consisting of all elements in $\mathfrak{G}$ of trace zero. Since $T(\Sigma x_i\lambda_i) = 2x_0 + x_1$, an element is in $\mathfrak{F}$ if and only if it may be written in the form $X = x_1(2\lambda_1 - 1) + x_2\lambda_2 + x_3\lambda_3$, where the x's are integers. The norm of the general element in $\mathfrak{F}$ is $N(X) = f(x_1, x_2, x_3)$ where f is the ternary quadratic form with the matrix

$$\Gamma = \begin{pmatrix} 2g_{11} - 1 & g_{12} & g_{13} \\ g_{21} & g'_{22} & \frac{1}{2}g_{23} \\ g_{31} & \frac{1}{2}g_{32} & g'_{33} \end{pmatrix}.$$

f is a classic form, i.e., the coefficients of all cross-product terms are even, if and only if d is even. Since $|M| = d^2$, it may be shown that $|\Gamma| = d^2/4$.

Suppose $\xi_0, \cdots, \xi_3$ form a normal basis of $\mathfrak{G}$. Then $2\xi_1 - 1, \xi_2, \xi_3$ form a basis of $\mathfrak{F}$ and therefore these elements are obtained from $2\lambda_1 - 1, \lambda_2, \lambda_3$ by an integral transformation of determinant ± 1. Hence the form $N[y_1(2\xi_1 - 1) +$

[7] Brandt, loc. cit., p. 12.

[8] Brandt, loc. cit., p. 10.

$y_2\xi_2 + y_3\xi_3]$ is equivalent to f. Hence the class $\mathfrak{C}_1$ of forms equivalent to f is independent of the particular normal basis employed and is uniquely determined by $\mathfrak{G}$.

More generally, suppose $\mathfrak{G}_1$ is an integral set in the same class as $\mathfrak{G}$. By Theorem 1, $\mathfrak{G}_1$ has a basis $\eta_0, \cdots, \eta_3$ such that $N(\Sigma x_i\eta_i) = \Phi$. Hence $\mathfrak{C}_1$ is uniquely determined by the class of integral sets containing $\mathfrak{G}$.

Suppose $d = 2d_1$ is even. It will be found that every second order minor of Γ is equal to one half or one fourth of one of the third order minors of M. Therefore by Theorem 1 every such second order minor is divisible by d_1. Since $| \Gamma | = d_1^2$, it follows that $\pm d_1$ is the g.c.d. of these second order minors. By the definition of d, Φ is positive or indefinite according as $d > 0$ or $d < 0$. Hence the same is true of f and the invariants of f are $\Omega = d_1$, $\Delta = 1$.[9] Since d_1 is odd and contains no square factor > 1, f is a properly primitive form.

If d is odd, it may be shown that the invariants of the improperly primitive form $2f$ are $\Omega = d$, $\Delta = 2$.

The class of forms containing f or $2f$, according as d is even or odd, will be said to correspond to the class of integral sets containing $\mathfrak{G}$.

Let F be a quadratic field which is imaginary if $d > 0$. It is known that $\mathfrak{A}$ contains a field equivalent to F if and only if no prime factor of d is the product of two distinct prime ideals in the set of all integral numbers of F.[10] If d is even, f is a properly primitive form and if d is odd, $2f$ is improperly primitive. Therefore, in either case, f represents a positive integer k, prime to $2d$.[11] Then $\mathfrak{F}$ contains an element η, such that $\eta^2 = -k$. Since $\mathfrak{A}$ contains the quadratic field $F(\eta)$, it follows that $-k$ is a quadratic non-residue of every odd prime factor, $q_1, q_2, \cdots, q_n$ of d. Hence if d is even, the characters of f are[12]

$$\chi_i = \left(\frac{k}{q_i}\right) = (-1)^{(q_i+1)/2} \qquad (i = 1, 2, \cdots, n). \tag{2}$$

If d is odd, the invariants of $2f$ are $\Omega = d$, $\Delta = 2$ and therefore by the last reference, the characters of $2f$ are

$$\chi_i = \left(\frac{2k}{q_i}\right) = (-1)^{(q_i+1)(q_i+3)/8} \qquad (i = 1, 2, \cdots, n). \tag{3}$$

We have then all but the last sentence of

THEOREM 2. *Let $\mathfrak{A}$ be a rational generalized quaternion algebra with the fundamental number d. For every class of integral sets in $\mathfrak{A}$ there is a uniquely deter-*

[9] For the definitions of these invariants, see Dickson's *Studies in the Theory of Numbers*, p. 10. This book will be referred to hereafter as *Studies*.

[10] This is a consequence of a more general theorem by Hasse, *Die Struktur der R. Brauerschen Algebrenklassengruppen über einem algebraischen Zahlkörper*, Mathematische Annalen, vol. 107 (1933), p. 731; Deuring, *Algebren*, p. 118. See also *The quadratic subfields of a generalized quaternion algebra*, this Journal, vol. 2, p. 681.

[11] Dickson, *Studies*, Theorems 6, 7, p. 8.

[12] Dickson, *Studies*, p. 52.

mined corresponding class $\mathfrak{C}$ *of non-negative ternary quadratic forms. If* $d = 2d_1$ *is even, the invariants of the forms in* $\mathfrak{C}$ *are* $\Omega = d_1$, $\Delta = 1$ *and their characters are the* χ_i *in* (2). *If* d *is odd, the forms in* $\mathfrak{C}$ *are improperly primitive and have the invariants* $\Omega = d$, $\Delta = 2$ *and the characters* χ_i *of* (3). *No class of forms corresponds to two classes of integral sets.*

To prove the last sentence of the theorem, we employ the following which will be proved in §6.

LEMMA 1. *Let* $\mathfrak{G}$, $\mathfrak{G}_1$ *be sets of integral elements in* $\mathfrak{A}$ *and let* $\mathfrak{F}$, $\mathfrak{F}_1$ *be the sets of elements in* $\mathfrak{G}$, $\mathfrak{G}_1$, *respectively, of trace zero. If* $\mathfrak{F} = \mathfrak{F}_1$, *then* $\mathfrak{G} = \mathfrak{G}_1$.

Let $\lambda_0, \cdots, \lambda_3$ and $\xi_0, \cdots, \xi_3$ be normal bases of the integral sets $\mathfrak{G}$ and $\mathfrak{G}_1$ respectively. Suppose the form $N[x_1(2\lambda_1 - 1) + x_2\lambda_2 + x_3\lambda_3] = f(x_1, x_2, x_3)$ is transformed into $N[y_1(2\xi_1 - 1) + y_2\xi_2 + y_3\xi_3] = f_1(y_1, y_2, y_3)$ by a transformation $x_i = \Sigma_j t_{ij} y_j$ $(i = 1, 2, 3)$, where the t's are integers, $|t_{ij}| = \pm 1$. If u is rational and $T(Z) = 0$, then $N(u + Z) = u^2 + N(Z)$. Hence $N(\Sigma y_i\xi_i) = y_0^2 + y_0y_1 + y_1^2/4 + f_1(y_1/2, y_2, y_3) = \Phi_1(y_0, \cdots, y_3)$. Let $\eta_0 = 1$ and let η_1, η_2, η_3 be defined by

$$
\begin{aligned}
2\eta_1 - 1 &= t_{11}(2\lambda_1 - 1) + t_{21}\lambda_2 + t_{31}\lambda_3, \\
\eta_2 &= t_{12}(2\lambda_1 - 1) + t_{22}\lambda_2 + t_{32}\lambda_3, \qquad (4)\\
\eta_3 &= t_{13}(2\lambda_1 - 1) + t_{23}\lambda_2 + t_{33}\lambda_3.
\end{aligned}
$$

Then $Y_1 = y_1(2\eta_1 - 1) + y_2\eta_2 + y_3\eta_3 = x_1(2\lambda_1 - 1) + x_2\lambda_2 + x_3\lambda_3$ and $N(Y_1) = f_1(y_1, y_2, y_3)$. Therefore if $Y = \Sigma y_i\eta_i$, then $N(Y) = \Phi_1(y_0, y_1, y_2, y_3) = N(\Sigma y_i\xi_i)$. Since the η's form a basis of $\mathfrak{A}$ and the ξ's form a basis of $\mathfrak{G}_1$, by Theorem 1, the η's form a normal basis of an integral set $\mathfrak{G}^1$ equivalent to $\mathfrak{G}_1$. But by (4) and Lemma 1, $\mathfrak{G}^1 = \mathfrak{G}$. Hence $\mathfrak{G}$ and $\mathfrak{G}_1$ belong to the same class. This completes the proof of Theorem 2.

If d is even, the forms in $\mathfrak{C}$ and also their reciprocals have odd determinants and therefore are properly primitive. Therefore by Theorem 2 and a known result,[13] we have the

COROLLARY. *If* $d < 0$, *any two integral sets in* $\mathfrak{A}$ *are equivalent.*

4. The correspondence between classes of integral sets and classes of forms. We shall now prove

THEOREM 3. *Let* $\mathfrak{A}$ *be a generalized quaternion algebra with the fundamental number* d. *If* d *is even, let* G *be the genus of forms with the invariants* $\Omega = d/2$, $\Delta = 1$ *and the characters* χ_i *of* (2). *If* d *is odd, let* G *be the genus of forms with the invariants* $\Omega = d$, $\Delta = 2$ *and the characters* χ_i *of* (3). *There is a one-to-one correspondence between the classes of integral sets in* $\mathfrak{A}$ *and the non-negative classes in* G *or the non-negative improperly primitive classes in* G, *according as* d *is even or odd.*

If $d < 0$, the theorem follows from Theorem 2 and the theorem cited in the proof of the corollary. Hence we shall assume that $d > 0$. By Theorem 2, it

[13] Dickson, *Studies*, Theorem 47, p. 54.

will be sufficient to show that every positive class or every improperly primitive positive class in G corresponds to a class of integral sets.

Suppose $d = 2d_1$ is even. Let $\mathfrak{C}$ be a class of positive forms in G. Let $\psi = \Sigma h_{ij}x_ix_j$ be the reciprocal of a form in $\mathfrak{C}$. We have seen that ψ is a properly primitive form and therefore it represents an integer prime to d.[14] We may then assume that h_{33} is prime to d. Let $f = \Sigma a_{ij}x_ix_j$ be the form in $\mathfrak{C}$ reciprocal to ψ, and let A_{ij} be the cofactor of a_{ij} in the matrix (a_{ij}). Then $a_{11}a_{22} - a_{12}^2 = A_{33} = d_1h$, where $h = h_{33}$ is prime to d. Consider the binary form $f(x_1, x_2, 0)$. Since d_1 contains no square factor > 1 and is prime to h, the g.c.d. of the coefficients of this form is prime to d_1. Therefore $f(x_1, x_2, 0)$ represents an integer prime to d_1. Hence we may also assume that a_{11} is prime to d_1. Let $h = ak^2$, $a_{11} = \alpha t^2$, where a, α contain no square factor > 1. In $f(x_1, x_2, x_3)$, set

$$\begin{aligned} x_1 &= x - \frac{a_{12}}{a_{11}}y + \frac{A_{13}}{A_{33}}z, \\ x_2 &= \qquad\quad y + \frac{A_{23}}{A_{33}}z, \\ x_3 &= \qquad\qquad\qquad\;\; z, \end{aligned}$$

and in the resulting form, replace x, y, z by x/t, $\alpha tz/k$, aky respectively. We obtain

$$\Psi(x, y, z) = \alpha x^2 + \beta y^2 + \alpha\beta z^2,$$

where $\beta = ad_1$.

Since h is prime to d_1, β contains no square factor > 1. Let $\alpha = \alpha_1\delta$, $\beta = \beta_1\delta$, where δ is the g.c.d. of α, β. Since a_{11} is prime to d_1, β_1 is divisible by d_1. Let B be the least positive divisor of β_1 such that $y^2 + \alpha \equiv 0 \pmod{\beta_1/B}$ has a solution. Since f represents $a_{11} = \alpha t^2$ and the characters of f are the χ_i of (2), it follows that $-\alpha$ is a quadratic non-residue of every prime factor of d_1. Hence d_1 divides B.

Let $\mathfrak{A}'$ be an algebra with the basis 1, i, j, ij, where $i^2 = -\alpha$, $j^2 = -\beta$, $ij = -ji$. The fundamental number d' of $\mathfrak{A}'$ is divisible by B.[15] Hence d' is divisible by d_1.

Since $\psi = \Sigma h_{ij}x_ix_j$ is properly primitive, after an interchange of variables necessary, we may assume that h_{11} is odd. Replace x_1 by $x_1 + \xi x_2 + \eta x_3$, where $\xi = 0$ or 1 according as h_{22} is even or odd and $\eta = 0$ or 1 according as h_{33} is even or odd. In the resulting form, $\psi_1 = \Sigma k_{ij}x_ix_j$, $k_{11} + 1 \equiv k_{22} \equiv k_{33} \equiv 0 \pmod 2$. Since $|k_{ij}| = d_1$ is odd, it follows that k_{23} is odd. Let $f_1 = \Sigma c_{ij}x_ix_j$ be the form in $\mathfrak{C}$ reciprocal to ψ_1. Then $c_{11} = k_{22}k_{33} - k_{23}^2$, $c_{11}c_{22} - c_{12}^2 = d_1k_{33}$, $c_{11}c_{33} - c_{13}^2 = d_1k_{22}$, and therefore

$$c_{11} \equiv -1 \pmod 4, \quad c_{22} \equiv c_{12}, \quad c_{33} \equiv c_{13} \pmod 2. \tag{5}$$

[14] Dickson, *Studies*, p. 14, Theorem 16.

[15] This Journal, vol. 1 (1935), p. 435.

Let g_{ij}, g'_{ij} be the integers defined by equating the matrix Γ of §3 to (c_{ij}) and let $g_{22} = 2g'_{22}$, $g_{33} = 2g'_{33}$. Consider the matrix $M = (g_{ij})$ of §2 with the elements g_{ij} as thus defined. Since $|c_{ij}| = d_1^2$, it may be shown that $|M| = 4d_1^2 = d^2$. Let $G_{00} = |g_{ij}|$, $(i, j = 1, 2, 3)$. We have $d^2 = 2G_{00} - (g_{22}g_{33} - g_{23}^2)$. Since $g_{22}g_{33} - g_{23}^2 = 4(c_{22}c_{33} - c_{23}^2) = 4d_1k_{11}$ and k_{11} is odd, it follows that G_{00} is divisible by $2d$. Noting that by (5), $g_{22} + 2g_{12}^2 \equiv g_{33} + 2g_{13}^2 \equiv 0 \pmod 4$, and employing the fact that every second order minor of (c_{ij}) is divisible by d_1, we may show that $M = (g_{ij})$ satisfies the remaining conditions of (c) in Theorem 1.

We have seen that f is transformed into Ψ by a non-singular transformation with rational coefficients. Hence $\Psi(y_1, y_2, y_3)$ is transformed into $f_1(x_1, x_2, x_3)$ by such a transformation, $y_i = \Sigma_j s_{ij}x_j$ $(i = 1, 2, 3)$. Let $\omega_1, \omega_2, \omega_3$ be elements of $\mathfrak{A}'$ defined by

$$
\begin{aligned}
2\omega_1 - 1 &= s_{11}i + s_{21}j + s_{31}ij, \\
\omega_2 &= s_{12}i + s_{22}j + s_{32}ij, \\
\omega_3 &= s_{13}i + s_{23}j + s_{33}ij.
\end{aligned}
$$

Then $\omega_0 = 1, \cdots, \omega_3$ form a basis of $\mathfrak{A}'$. We have $X_1 = x_1(2\omega_1 - 1) + x_2\omega_2 + x_3\omega_3 = y_1i + y_2j + y_3ij$ and therefore $N(X_1) = f_1(x_1, x_2, x_3)$. Since $T(X_1) = 0$, it follows that if $X = \Sigma x_i\omega_i$, then $N(X) = x_0^2 + x_0x_1 + x_1^2/4 + f_1(x_1/2, x_2, x_3) = \Phi(x_0, x_1, x_2, x_3)$, where M is the matrix of 2Φ.

We may set $1 = i_0$, $i = \alpha^{\frac{1}{2}}i_1$, $j = \beta^{\frac{1}{2}}i_2$, $ij = (\alpha\beta)^{\frac{1}{2}}i_3$, where the i_j are the hamiltonian quaternion units. Then $\omega_n = \Sigma_k t_{kn}i_k$, where the t's are real. Since the determinant of $N(\Sigma x_i\omega_i)$ is $d^2/16$, it follows that, after changing the sign of ω_3 if necessary, $d = 4\,|t_{kn}|$.

Feuter considered an integral set in a quaternion algebra with the fundamental number d.[16] If $\omega_0 = 1, \cdots, \omega_3$ form a basis of this set, he determined explicitly the integers c_{ijk} defined by $\omega_i\omega_j = \Sigma_k c_{ijk}\omega_k$ in terms of the coefficients of the form $N(\Sigma x_i\omega_i) = \frac{1}{2}\Sigma g_{ij}x_ix_j$ and the $\mathfrak{g}$'s defined by certain relations between the ω's. Since his ω's form a basis of an integral set, it follows from his definitions of the $\mathfrak{g}$'s that they are integers.

Consider Fueter's argument, beginning on p. 651 and leading to equations (11) on p. 654, as applied to the present ω's. The argument holds without modification except that the definitions of the $\mathfrak{g}$'s now imply only that they are rational. However, since they satisfy his equations (7), p. 653, it follows that every $d\mathfrak{g}_{ij}$ is equal to the cofactor of one of the elements of M. Since M satisfies (c) of Theorem 1, it follows that every $\mathfrak{g}_{ij}$ is an integer. Hence by Fueter's equations (11), p. 654, $\omega_i\omega_j = \Sigma c_{ijk}\omega_k$, where each c is either an integer or half an odd integer. By a result due to Brandt, the c's are all integral.[17]

Let $\mathfrak{G}$ be the set of all elements $X = \Sigma x_i\omega_i$ with integral x's. By definition

[16] *Zur Theorie der Brandtschen Quaternionenalgebren*, Mathematische Annalen, vol. 110 (1934–5), pp. 651–4.

[17] *Der Kompositionsbegriff bei den quaternären quadratischen Formen*, Mathematische Annalen, vol. 91 (1924), p. 303.

of the ω's, $T(X) = 2x_0 + x_1$. Since $N(X) = \Phi(x_0, \cdots, x_3)$, it follows that $\mathfrak{G}$ has the property R used in our definition of an integral set. By the preceding paragraph, $\mathfrak{G}$ has the property C. It obviously has the property U. Let $\mathfrak{G}'$ be an integral set in $\mathfrak{A}'$ containing $\mathfrak{G}$, i.e., a maximal set with the properties R, C, U and containing $\mathfrak{G}$. The norm of the general element in $\mathfrak{G}'$ is a form Φ' and by Theorem 1, the determinant of $2\Phi'$ is d'^2. Since Φ' is transformed into Φ by an integral transformation, it follows that d' divides d. But we have seen that d_1 divides d'. Hence $d' = d_1$ or $d' = d$. Since every positive fundamental number is the product of an odd number of prime factors, it follows that $d' = d$ and hence $\mathfrak{A}'$ is equivalent to $\mathfrak{A}$.[18] We may therefore identify $\mathfrak{A}'$ with $\mathfrak{A}$. Then $\mathfrak{G}' = \mathfrak{G}$ is an integral set in $\mathfrak{A}$. Since $N[x_1(2\omega_1 - 1) + x_2\omega_2 + x_3\omega_3] = f_1(x_1, x_2, x_3)$, it follows that $\mathfrak{C}$ is the class of forms corresponding to the class of integral sets containing $\mathfrak{G}$. This proves the theorem for the case where d is even.

Suppose d is odd. Let $\mathfrak{C}$ be a class of positive improperly primitive forms in G and let $2f = \Sigma a_{ij}x_ix_j$ be a form in $\mathfrak{C}$. Since $| a_{ij} | = 2d^2$, one of the a_{ij}, say a_{23}, is odd. If a_{12} is odd, replace x_3 by $x_1 + x_3$ and if a_{13} is odd, replace x_2 by $x_1 + x_2$. We may then assume $a_{12} \equiv a_{13} \equiv a_{23} + 1 \equiv 0 \pmod 2$. Every term in the expansion of $| a_{ij} |$ is divisible by 8 except $-a_{11}a_{23}^2$. Hence $a_{11} \equiv -2 \pmod 8$. Let g_{ij}, g'_{ij} be the integers defined by equating the matrix Γ of §3 to the matrix $(a_{ij}/2)$ of f, let $g_{22} = 2g'_{22}$, $g_{33} = 2g'_{33}$ and let M be the matrix of §2 with the elements g_{ij} as thus defined. Since the invariants of $2f$ are $\Omega = d$, $\Delta = 2$, it may be shown that $M = (g_{ij})$ satisfies all the conditions of Theorem 1. We may then show, in the same manner as for the case where d is even, that there is an integral set $\mathfrak{G}$ in $\mathfrak{A}$ with a normal basis $\omega_0, \cdots, \omega_3$ such that $N(\Sigma x_i\omega_i) = \Phi(x_0, \cdots, x_3)$, where M is the matrix of 2Φ and $N[x_1(2\omega_1 - 1) + x_2\omega_2 + x_3\omega_3] = f(x_1, x_2, x_3)$. Then $\mathfrak{C}$ is the class of forms corresponding to the class of integral sets containing $\mathfrak{G}$ and the theorem is proved.

5. **A basis of an integral set.** Let F be a quadratic field with the discriminant $-\tau$, where τ is a prime not a divisor of d. By a result previously referred to, $\mathfrak{A}$ contains a field equivalent to F if and only if no prime factor of d is the product of two prime ideals in F. This condition is satisfied if and only if $-\tau$ is a quadratic non-residue of every odd prime factor of d and $\tau \equiv 3 \pmod 8$ if d is even. It is known that for every such field, called a canonical splitting field, $\mathfrak{A}$ has a basis $1, i, j, ij$, where $i^2 = -\tau$, $j^2 = -d$, $ij = -ji$.[19] Such a basis of $\mathfrak{A}$ will be called a canonical basis.

Let $\mathfrak{G}$ be an integral set in $\mathfrak{A}$ and let F be a canonical splitting field of $\mathfrak{A}$ with the discriminant $-\tau$. The intersection of $\mathfrak{G}$ and F is a ring with a basis 1, $c\omega$, where 1, ω form a basis of all the integral elements of F and c is a positive integer. Hull determined a basis v_0, v_1, v_2, v_3 of $\mathfrak{G}$ such that $v_0 = 1$, $v_1 = c\omega$.[20]

[18] Brandt, loc. cit., pp. 12, 13.

[19] This Journal, vol. 1 (1935), p. 435.

[20] Transactions of the American Mathematical Society, vol. 40 (1936), p. 8.

c is not in general equal to unity. We shall show that there is an infinitude of canonical splitting fields such that for each of them, $c = 1$.[21] Choosing F as such a field and setting $c = 1$ in the v_i, we obtain from them a relatively simple normal basis of $\mathfrak{G}$.

Let $\omega_0, \cdots, \omega_3$ be a normal basis of $\mathfrak{G}$ and let $f(x_1, x_2, x_3)$ be as defined in §3. We shall first prove

LEMMA 2. *$f(x_1, 2x_2, 2x_3)$ represents an infinitude of primes.*

Let $d' = d/2$ and $f'(x_1, x_2, x_3) = f(x_1, x_2, x_3)$ or let $d' = d, f'(x_1, x_2, x_3) = f(x_1, x_2, 2x_3)$ according as d is even or odd. If d is even, we have seen that the invariants of f' are $\Omega = d', \Delta = 1$ and the reciprocal ψ' of f' is properly primitive. It may be shown that f' and ψ' have these properties if d is odd.

For properly chosen integers C_i, $\psi'(C_1, C_2, C_3) = C > 0$, where C is prime to d'.[22] After adding properly chosen multiples of d' to the C_i, we may assume $C_1 \equiv 0$, $C_2 \equiv C_3 \equiv \pm 1 \pmod 4$. Finally, without disturbing the preceding conditions, we may assume that the g.c.d. of the C_i is unity. Let $a_1 = (C_2 + C_3)/\delta$, $a_2 = a_3 = -C_1/\delta$, where δ is the g.c.d. of C_1 and $C_2 + C_3$. Then a_1 is odd, a_2 and a_3 are even, the g.c.d. of the a_i is unity and $\Sigma a_i C_i = 0$. We may then determine integers b_i, such that[23]

$$\begin{vmatrix} a_2 & b_2 \\ a_3 & b_3 \end{vmatrix} = C_1, \qquad \begin{vmatrix} a_1 & b_1 \\ a_3 & b_3 \end{vmatrix} = -C_2, \qquad \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} = C_3.$$

In $f'(x_1, x_2, x_3)$, set $x_i = a_i y_1 + b_i y_2$ $(i = 1, 2, 3)$. We obtain a primitive form $\phi(y_1, y_2) = ay_1^2 + 2ty_1y_2 + by_2^2$, where $ab - t^2 = Cd'$.[24] We have $a = f'(a_1, a_2, a_3)$. Since the leading coefficient of f' is $\equiv 3 \pmod 4$ and a_1 is odd, a_2 and a_3 are even, it follows that $a \equiv 3 \pmod 4$. Then $\phi(y_1, 2y_2)$ is a properly primitive form whose discriminant is not a perfect square. Hence it represents an infinitude of primes. Since a_2, a_3 are even, the lemma follows.

Let τ be one of the infinitude of primes represented by $f(x_1, 2x_2, 2x_3)$ and not a divisor of d. Then there is an element $i = x_1(2\omega_1 - 1) + 2x_2\omega_2 + 2x_3\omega_3$ in $\mathfrak{F}$ such that $i^2 = -\tau$. Since x_1 is odd, $\mathfrak{G}$ contains $\lambda_1 = (1 + i)/2$. The field $F(i)$ is a canonical splitting field of $\mathfrak{A}$ and $\mathfrak{A}$ has a basis $1, i, j, ij$, where $j = -d$, $ij = -ji$.

Identify Hull's Q, $\mathfrak{M}$, p, u, P, a_0[25] with our $\mathfrak{A}$, $\mathfrak{G}$, τ, j, $F(i)$, $-d$ respectively. Then $\mathfrak{G}$ has a basis in the form below,[26] if we note that $\mathfrak{G}$ contains λ_1, and hence Hull's $c = 1$.

$$v_0 = 1, \quad v_1 = (-\tau + i)/2, \quad v_2 = g(\lambda + j), \quad v_3 = g(\lambda + j)\left[\frac{\tau(2h - 1) + i}{2\tau g^2}\right].$$

[21] Cf. Hull, loc. cit., p. 11, lines 9–13.

[22] Dickson, *Studies*, p. 14, Theorem 16.

[23] Dickson, *Studies*, p. 11, Theorem 9.

[24] Dickson, *Studies*, Theorem 27, p. 25 and Theorem 37, p. 32.

[25] Loc. cit., Theorem 6, p. 7.

[26] Hull, loc. cit., p. 8.

Employing Hull's equations (16) and (17), p. 7, loc. cit., and setting $b = 1 - 2h$, we find

$$v_3 = k_2(1 + bi)/2 + \frac{k_1 g_3}{\tau} i - \frac{b}{2g} j - \frac{1}{2\tau g} ij.$$

Since b is odd and $g\lambda$ is an integral element in $F(i)$, it follows that $\mathfrak{G}$ has a basis $\lambda_0, \cdots, \lambda_3$ as below. $N(\lambda_3)$ is an integer and therefore the congruences (5) are satisfied.

To prove the last sentence of the following theorem, let $\mathfrak{G}$ be the set of all elements $\Sigma x_i \lambda_i$ with integral x's. It may be verified that $\mathfrak{G}$ has the properties R, C, U used in our definition of an integral set. Since the form $\Phi = N(\Sigma x_i \lambda_i)$ is obtained from $N(x + yi + zj + wij)$ by a transformation of determinant $(4\tau)^{-1}$, it follows that the determinant of 2Φ is d^2. Hence $\mathfrak{G}$ is maximal. We have then

THEOREM 4. *Let $\mathfrak{G}$ be an integral set in $\mathfrak{A}$. There is an infinitude of primes τ such that for every τ there is an element i in $\mathfrak{G}$ such that* (a) $i^2 = -\tau$, (b) *the field $F(i)$ is a canonical splitting field of $\mathfrak{A}$*, (c) *$\mathfrak{G}$ contains $(1 + i)/2$. For every such i, there is a canonical basis $1, i, j, ij$ of $\mathfrak{A}$ and $\mathfrak{G}$ has a basis in the form*

$$(4) \qquad \lambda_0 = 1, \qquad \lambda_1 = (1 + i)/2, \qquad \lambda_2 = gj, \qquad \lambda_3 = \frac{a}{\tau} i + \frac{b}{2g} j + \frac{1}{2\tau g} ij,$$

where $a, b, g > 0$ are integers such that

$$(5) \qquad 4a^2g^2 + d \equiv 0 \pmod{\tau}, \qquad \tau b^2 + 1 \equiv 0 \pmod{4g^2}.$$

Conversely, if $1, i, j, ij$ form a canonical basis of $\mathfrak{A}$, $i^2 = -\tau$, and if a, b, g are integers which satisfy (5), *then $\lambda_0, \cdots, \lambda_3$ of* (4) *form a normal basis of an integral set.*

We shall show by an example that a canonical splitting field $F(i)$ cannot always be chosen so that an integral set has a basis in the form (4) with $g = 1$. Let $\mathfrak{A}$ be the algebra with a basis $1, I, J, IJ$, where $I^2 = -83$, $J^2 = -10$, $IJ = -JI$. The fundamental number of $\mathfrak{A}$ is $d = 2 \cdot 5 \cdot 83$.[27] It may be verified that the following elements form a basis of an integral set $\mathfrak{G}$ in $\mathfrak{A}$.

$$\omega_0 = 1, \qquad \omega_1 = \tfrac{1}{2}(1 + I), \qquad \omega_2 = 3J, \qquad \omega_3 = \tfrac{1}{6}(-5 + I)J.$$

It may also be verified that $\mathfrak{G}$ contains no element of trace zero and norm d. Hence there is no canonical splitting field $F(i)$ such that $\mathfrak{G}$ has a basis in the form (4) with $g = 1$.

6. **Proof of Lemma** 1. Let $\lambda_0, \cdots, \lambda_3$ of (4) be a basis of $\mathfrak{G}$. Since $2\lambda_1 - 1 = i$ is in $\mathfrak{G}_1$, the intersection of $\mathfrak{G}_1$ with the set of all integral elements of $F(i)$ is a ring with the conductor $c = 1$ or $c = 2$. If $c = 1$, $\mathfrak{G}_1$ contains λ_1 and, since $T(\lambda_2) = T(\lambda_3) = 0$, $\mathfrak{G}_1$ also contains λ_2, λ_3. Hence $\mathfrak{G}_1$ contains $\mathfrak{G}$ and the lemma follows. We shall show that the case $c = 2$ cannot occur.

[27] This Journal, vol. 1 (1935), p. 435.

Suppose $c = 2$. By Hull's result,[28] $\mathfrak{G}_1$ has a basis in the form

$$v_0 = 1, \qquad v_1 = i, \qquad v_2 = (\lambda + j)g_1/2, \qquad v_3 = (\lambda + j)\left(\frac{\tau h - \tau + i}{2g_1}\right),$$

where λ is a certain element in $F(i)$ and g_1 is a positive integer. If $\zeta_0, \cdots, \zeta_3$ is a normal basis of $\mathfrak{G}_1$, the trace of each of the elements $2\zeta_1 - 1, \zeta_2, \zeta_3$ is zero. It follows that the double of every element of $\mathfrak{G}_1$ is in $\mathfrak{G}$. In particular, $2v_2$ is in $\mathfrak{G}$. Hence g divides g_1. Since λ_2 is in $\mathfrak{G}_1$, g_1 divides $2g$. But by Hull's Theorem 9 (loc. cit., p. 10), g_1 is odd. Hence $g = g_1$. By Hull's Theorems 7, 8, the lemma on p. 9 and (16) on p. 7, d and h are odd. Then in his (17), $k_1 \equiv k_2 + 1 \pmod 2$ and we find

$$v_2 - \frac{k_1}{2} = zi + \frac{g}{2}j, \qquad v_3 - \frac{k_2}{2} + \lambda_3 = \left(w + \frac{a}{\tau}\right)i + \frac{h - 1 + b}{2g}j,$$

where z, w are certain rational numbers. Since k_1 or k_2 is even and $T(\lambda_3) = 0$, the right member of one of these equations is in $\mathfrak{G}$ and hence is expressible as a linear function of $\lambda_0, \lambda_1, \lambda_2$ with integral coefficients. But this is obviously impossible in the first case and, since h and b are odd, it is impossible in the second case. Therefore $c \neq 2$ and the lemma follows.

7. **Certain integral sets.** Let I, J be elements of $\mathfrak{A}$ such that
(a) $1, I, J, IJ$ form a basis of $\mathfrak{A}$;
(b) $I^2 = -\alpha, J^2 = -\beta, IJ = -JI$, where α, β are integers, neither divisible by a square > 1 and $\alpha \equiv \beta \pmod 2$.

All the integral sets, finite in number, which contain I and J were determined by Darkow[29] and the writer,[30] the former treating the case where α and β are even. Of course, α and β are not uniquely determined by $\mathfrak{A}$. Let $\mathfrak{G}$ be an arbitrarily chosen integral set in $\mathfrak{A}$. The question arises as to whether or not $\mathfrak{G}$ may be obtained from the above mentioned results by proper choice of I and J, i.e., whether or not $\mathfrak{G}$ contains elements I and J which satisfy the conditions (a) and (b). We shall show that it does contain such elements.

Let $\lambda_0, \cdots, \lambda_3$ of Theorem 4 be a basis of $\mathfrak{G}$. $\mathfrak{G}$ contains $i = 2\lambda_1 - 1$,

$$J = -axi + y\lambda_2 + \tau x\lambda_3 = (\tau bx + 2g^2y + xi)j/2g$$

and iJ, where x and y are rational integers to be determined later. We have $J^2 = -d\phi(x, y)$, where $\phi(x, y) = \tau cx^2 + \tau bxy + g^2y^2$, $c = (\tau b^2 + 1)/4g^2$. The

[28] Loc. cit., p. 8.

[29] *Determination of a basis for the integral elements of certain generalized quaternion algebras*, Annals of Mathematics, vol. 28 (1926–27), pp. 263–270.

[30] *Arithmetics of generalized quaternion algebras*, American Journal of Mathematics, vol. 48 (1926), pp. 57–63.

discriminant of ϕ is $-\tau$ and hence it represents an infinitude of primes. Therefore, for properly chosen integers x and y, $J^2 = -dp$, where p is a prime not a divisor of $2d\tau$. Then dp has no square factor > 1 and is prime to τ. It may be shown that $1, i, J, iJ$ are linearly independent. Hence they form a basis of $\mathfrak{A}$; furthermore, $iJ = -Ji$.

Let $I = i$ or $I = iJ$ according as d is odd or even. Then I and J satisfy all the conditions (a) and (b). Since $\mathfrak{G}$ contains I and J, it follows that it is one of the sets obtained by properly specializing the above mentioned results by Darkow or by the writer, according as d is even or odd.

UNIVERSITY OF KENTUCKY.

THE ERGODIC FUNCTION OF BIRKHOFF

By Monroe H. Martin

Introduction. In his New Orleans lecture, Professor Birkhoff[1] introduced the concept of the ergodic function $T(\epsilon)$ as the least time T which elapses before the point P of some motion can come within a distance ϵ of every point of the phase space. He is led to conjecture that in the general closed recurrent case possessing no stable periodic motions the ergodic function is of the order of $\epsilon^{-(n-1)}$, n being the number of dimensions of the phase space. The purpose of this paper is to consider the closed, transitive dynamical systems provided by the geodesics on topologically closed surfaces of constant negative curvature. The principal result is the establishment of upper bounds for the ergodic functions of these dynamical systems.

The metric chosen for the phase space is patterned after that used by Morse.[2] The "time" T along a motion is taken as the "H-length" measured along a geodesic on the surface (see §9 of the present paper).

If ϵ be chosen in compliance with (67), (68), we find that the ergodic function $T(\epsilon)$ satisfies an equality of the form

$$T(\epsilon) < \epsilon^{-\omega}\left[A \log \frac{B}{\epsilon} + C\right],$$

where A, B, C and ω depend on the genus p of the surface as set forth in (77), (78), (79) and (80). All of these constants tend to $+\infty$ for $p \to +\infty$, the constant ω, in particular, being bounded by the inequalities (81), so that $\omega > 4 > 2$ for $p = 2, 3, \cdots$. Since $n - 1 = 2$ in the case under consideration, the order of $T(\epsilon)$ as conjectured by Birkhoff lies well below that of the upper bound we have found.

It may not be amiss to point out that the upper bound we have secured can doubtless be sharpened considerably, for in deriving it we compute an upper bound for the magnitude of the interval of time during which a point P of a certain motion comes within a distance ϵ of every point of the phase space at *least once*, the point P perchance coming *repeatedly* within a distance ϵ of some or all of the points of the phase space. In addition, the upper bound, $(4p)^{2m}$, given in §10 for the number of sets S_M^* can in all probability be improved upon, with a resultant improvement in the upper bound for the ergodic function.

Received August 23, 1935. This paper had its inception while the author was National Research Fellow at Harvard University in 1933.

[1] G. D. Birkhoff, Bull. Amer. Math. Soc., vol. 38 (1932), pp. 375–377.

[2] M. Morse, Jour. de Math., vol. 14 (1935), pp. 52–53. This paper will be referred to as Morse I.

1. **Preliminary notions.** Let (x, y) denote ordinary Cartesian coördinates in the plane. Let Ψ denote the unit circle $x^2 + y^2 = 1$, and let the interior of Ψ, provided with the metric

$$\text{(H)} \qquad ds^2 = 4(1 - x^2 - y^2)^{-2}(dx^2 + dy^2),$$

be denoted by Φ. The Gaussian curvature of this metric equals -1, so that Φ may be regarded as the hyperbolic plane of non-euclidean geometry. Linear fractional transformations of hyperbolic type on the complex variable $x + iy$ which take the interior of Ψ into itself leave (H) invariant, and are called *H-transformations* (hyperbolic transformations). The geodesics on Φ are circular arcs orthogonal to Ψ and ending upon it. They are called *H-lines*. Through two points on Φ only one H-line can be drawn. Consequently if two H-lines intersect in more than one point, they coincide. Two H-lines which have no common point either on Φ or on Ψ are said to be *non-intersecting*. A segment of an H-line which has both end-points A, B in Φ is an *H-line segment*, and is denoted by AB; a segment which has one end-point in Φ and one on Ψ is an *H-ray*. The H-line of which a given H-line segment is a segment is, at times, conveniently referred to as the *H-line segment produced*. If A, B are two distinct points on $\Psi(\Phi)$, there is exactly one H-line (H-line segment) joining them. On directing this H-line (H-line segment) from A to B, we obtain a *directed H-line* $\overrightarrow{AB}$ (*directed H-line segment* $\overrightarrow{AB}$) for which A is the *α-end-point* and B the *ω-end-point*. If an H-line (H-line segment) and a segment of itself are directed in the same sense, the (latter) directed H-line segment is a *segment* of the (former) directed H-line (directed H-line segment). A directed H-line (directed H-line segment) *intersects* a set of directed H-line segments if one of the directed H-line segments in the set is a segment of the directed H-line (directed H-line segment). A set S_1 of directed H-line segments is a *subset of a set* S of directed H-line segments (directed H-lines) if every member of S_1 is a segment of some member of S.

The length of an H-line segment AB calculated in the metric (H) is its *H-length*, and is denoted by $|AB|$. The *H-distance between two points* A, B *of* Φ is then defined to be $|AB|$. When an H-line segment is designated by a single letter, say n, its H-length is designated[3] by $|n|$. The *H-mid-point* of an H-line segment is defined to be that point of the H-line segment which divides it into two H-line segments of equal H-lengths. Angular magnitudes measured in the metric (H) equal the ordinary Euclidean magnitudes. No distinction will therefore be made between the two. If A is a point not lying on a given H-line l and B the point of l such that AB meets l at right angles, AB is the *H-perpendicular* let fall from A to l. $|AB|$ is the *H-distance between* A *and* l and is, as a matter of fact, the minimum H-distance between A and the points of l. If two H-lines l, m are non-intersecting, they possess a unique mutual H-

[3] At times, we also employ the symbol $|\ |$ in its customary usage, that of denoting absolute values. No confusion need arise, since the proper interpretation will be clear from the context.

perpendicular, the H-length of which is denoted by $|lm|$, and is called the *H-distance between l and m.* $|lm|$ is then the minimum H-distance between the points of l and m.

If A, B, C are three non-coincident points of Φ such that two of the H-line segments AB, BC, CA meet at right angles, say BC and CA, the following relations[4] hold:

$$(1) \qquad \sin A = \frac{\sinh|BC|}{\sinh|AB|}, \qquad \cos A = \cosh|BC| \frac{\sinh|AC|}{\sinh|AB|},$$
$$\cosh|AB| = \cosh|AC| \cosh|BC|,$$

where A denotes the angle CAB. If A, B, C are three arbitrary points of Φ, the following "triangle" inequality holds:

$$|AB| \leqq |BC| + |CA|.$$

2. **Gaussian geodesic, and geodesic polar coördinates.** Let n be a given directed H-line segment. By a *Gaussian geodesic coördinate system* $[u_n, v_n]$ on Φ, we mean a coördinate system on Φ in which the coördinate lines v_n = const. are H-lines orthogonal to the directed H-line of which n is a segment, and the coördinate lines u_n = const. are the orthogonal trajectories of the coördinate lines v_n = const. The origin $[0, 0]$ is taken as the α-end-point of n. The positive sense on $u_n = 0$ is chosen to coincide with that on n and the sense on v_n = const. is then affixed in the usual way, i.e., so that an observer moving in the positive sense along $u_n = 0$ finds the positive sense of v_n = const. directed to his right. A unique coördinate system $[u_n, v_n]$ is thereby *associated* with any directed H-line segment n. The coördinate lines u_n = const. are circular arcs in Φ which connect the end-points of $u_n = 0$. They cut off equal H-lengths upon the coördinate lines v_n = const. If we take u_n as the H-length of arc measured along v_n = const. from $u_n = 0$, and v_n as the H-length of arc measured along $u_n = 0$ from $[0, 0]$, the differential form (H) becomes[5]

$$\text{(G)} \qquad ds^2 = du_n^2 + \cosh^2 u_n dv_n^2$$

in the $[u_n, v_n]$ coördinate system. Hence the H-length κ of the segment of $u_n = k$ comprehended between $v_n = 0$ and $v_n = |n|$ is given by

$$(2) \qquad \kappa = \int_0^{|n|} \cosh k dv_n = |n| \cosh k.$$

Let P, P' be two points of $u_n = a$. Denote by σ the H-length of the segment PP' of $u_n = a$. Let α, α' denote the directions (measured in the usual way in

[4] See, for example, H. S. Carslaw, *The Elements of Non-Euclidean Plane Geometry and Trigonometry*, London, 1916, p. 109.

[5] See, for example, W. Blashke, *Vorlesungen über Differentialgeometrie*, vol. 1, 1930 pp. 155–156.

the x, y-plane) of the coördinate lines v_n = const. at P, P' respectively. We proceed to calculate an upper bound for $|\alpha - \alpha'|$ when P, P' both lie in the region

$$-\delta \leqq u_n \leqq \delta, \qquad\qquad 0 \leqq v_n \leqq |n|$$

of Φ. Since the coördinate lines v_n = const. are orthogonal to $u_n = a$, it is sufficient to obtain an upper bound for the difference $|\beta - \beta'|$ in the directions β, β' of $u_n = a$ at P and P' respectively.

Let us assume (with no loss of generality) that $u_n = 0$ is drawn orthogonal to the positive x-axis of the x, y-coördinate system at a point which lies at an H-distance ξ from the origin of the x, y-coördinate system. The equation of $u_n = a$ in x, y-coördinates is then

$$x^2 + y^2 + \frac{2 \cosh \xi}{\sinh a - \sinh \xi}\, x + \frac{\sinh \xi + \sinh a}{\sinh \xi - \sinh a} = 0.$$

Letting K denote the ordinary Euclidean curvature of this circle, we find

$$|K| = \frac{|\sinh a - \sinh \xi|}{\cosh a}.$$

Now

$$\left|\frac{d\beta}{ds}\right| = \left|\frac{d\beta}{ds'}\right| \left|\frac{ds'}{ds}\right| = |K| \left|\frac{ds'}{ds}\right|,$$

where s, s' denote the H-length and Euclidean length of arc respectively, measured along $u_n = a$. From (H), we have $|ds'/ds| < \frac{1}{2}$, so that $|d\beta/ds| < \frac{1}{2}|K|$. Therefore

$$|\beta - \beta'| \leqq \int \left|\frac{d\beta}{ds}\right| ds < \tfrac{1}{2}|K|\sigma,$$

and hence we have from (2), since $\sigma \leqq |n| \cosh a$,

$$|\beta - \beta'| < \frac{|n|}{2} |\sinh a - \sinh \xi|,$$

$|K|$ being replaced by its value given above. Since $-\delta \leqq a \leqq \delta$, $\xi \geqq 0$ and $|\alpha - \alpha'| = |\beta - \beta'|$, we find that

(3) $$|\alpha - \alpha'| < \frac{|n|}{2} (\sinh \delta + \sinh \xi),$$

which is the required upper bound.

The set of H-rays concurring at a point O of Φ combined with their orthogonal trajectories form a *geodesic polar coördinate system* $[r, \psi]$ on Φ. r is the H-length

measured along an H-ray from O, and ψ is the angle at O measured from a fixed direction through O. In these coördinates (H) becomes[6]

$$ds^2 = dr^2 + \sinh r^2 d\psi^2. \tag{P}$$

The point O is called the *pole*.

A coördinate line r = const. will be called an *H-circle* and O is its *H-center*. The H-line segments connecting the points of an H-circle to its H-center are its *H-radii*. The two halves into which an H-circle is divided by an H-line through its H-center are *H-semicircles*. An H-circle is an ordinary Euclidean circle. However, its Euclidean center coincides with its H-center only when the H-center lies at the center of Ψ.

3. **Preliminary lemmas.** In this section we prove a number of elementary lemmas.

LEMMA 1. *If BC and AD are two H-line segments perpendicular to the H-line segment CD at its end-points, and if the H-line segment AB meets BC at right angles,*[7] *then*

$$\sinh |AB| = \sinh |CD| \cosh |AD|, \tag{4}$$

$$\cosh |CD| = \sin \angle BAD \cosh |AB|. \tag{5}$$

Draw the H-line segment AC. (4) follows from (1) and the relation $\cos \angle ACD = \sin \angle ACB$.

From (1) we have $\cosh |AC| = \cosh |AB| \cosh |BC| = \cosh |AD| \cosh |CD|$, and in addition, since $\sin \angle ACD = \cos \angle ACB$, there results $\sinh |AD| = \cosh |AB| \sinh |BC|$. (5) follows when these relations are used together with (1) and (4) to simplify the expression for $\sin \angle BAD$, obtained when $\sin \angle BAD$ is expressed in terms of sines and cosines of $\angle CAD$ and $\angle CAB$.

Let n be a given directed H-line segment and $[u_n, v_n]$ be the associated Gaussian geodesic coördinate system. λ_1, ρ denote positive numbers and C_1', C_1'' are two H-semicircles of H-radius ρ drawn in the region of Φ for which $v_n \geqq 0$ about the points $[\pm(\lambda_1 + \rho), 0]$ as H-centers.

LEMMA 2. *If C is an H-semicircle of H-radius ρ drawn in the region $v_n \geqq 0$ of Φ about a point on the segment $-(\lambda_1 + \rho) \leqq u_n \leqq \lambda_1 + \rho$ of $v_n = 0$ as H-center, it is intersected by any H-line l which intersects both C_1', C_1''.*

First, suppose $[0, 0]$ coincides with the center of Ψ. The coördinate lines $u_n = 0$, $v_n = 0$ are then perpendicular diameters of Ψ. Consider the H-line t_1

[6] See, for example, L. P. Eisenhart, *A Treatise on the Differential Geometry of Curves and Surfaces*, New York, 1909, pp. 207–209. Formula (P) above is not given explicitly here, but may be readily obtained from the formulas to be found on the pages indicated. As a matter of fact, (P) is a well-known formula for the arc length in non-Euclidean geometry.

[7] The H-line segments BC, AD, CD, AB form a figure which is sometimes called the tri-rectangular quadrilateral. See, for example, D. M. Y. Sommerville, *Non-Euclidean Geometry*, Chicago, 1919, p. 70.

drawn tangent to both C_1', C_1''. In order to prove the lemma, it is sufficient to prove it when l coincides with t_1. The H-distance from a point of the segment $-(\lambda_1 + \rho) \leqq u_n \leqq \lambda_1 + \rho$ of $v_n = 0$ to t_1 is seen from (4) to be a monotone increasing function of its H-distance from $[0, 0]$, and as such takes its greatest value, namely ρ, at the end-points of this segment. Since the H-center of C lies on this segment, C meets t_1.

When $[0, 0]$ does not coincide with the center of Ψ, an H-transformation carrying $[0, 0]$ into the center of Ψ reduces this case to the one above.

LEMMA 3. *If C', C'' are two H-semicircles of H-radii ρ drawn in the region $v_n \geqq 0$ of Φ about points on the segment $-(\lambda_1 + \rho) \leqq u_n \leqq \lambda_1 + \rho$ of $v_n = 0$ as H-centers, the set L of H-lines intersecting C', C'' contains the set L_1 of H-lines intersecting two arbitrarily chosen H-radii of C_1', C_1''.*

An H-line of L_1 intersects both C_1', C_1''. According to Lemma 1, it intersects both C', C'' and consequently belongs to L.

LEMMA 4. *The interior of the region of Φ enclosed by $v_n = 0$, two arbitrary radii of C_1', C_1'', and the H-line segment connecting their end-points on C_1', C_1'' cannot contain an H-circle of H-radius $\rho/2$.*

The proof of this lemma is simple and is omitted.

LEMMA 5. *If C', C'' are two H-semicircles of H-radius ρ drawn in the region $v_n \geqq 0$ of Φ about points $[\pm(\lambda + \rho), 0]$ $(\lambda > 0)$ as H-centers, and if n is such that $v_n = |n|$ is tangent to both C', C'', we have $|n|$ bounded by the inequalities*

$$\rho e^{-(\lambda+\rho)} < |n| < 2e^{-(\lambda+\rho)} \sinh \rho. \tag{6}$$

Fixing our attention on one of the H-semicircles, we let A denote its H-center, B its point of tangency with $v_n = |n|$, C the point $[0, |n|]$, and D the point $[0, 0]$. According to (4), we have

$$\sinh \rho = \sinh |n| \cosh (\lambda + \rho),$$

since $|AB| = \rho$, $|CD| = |n|$, $|AD| = \lambda + \rho$, and the upper bound in (6) follows at once from this equation, since

$$\sinh |n| > |n|, \qquad \cosh (\lambda + \rho) > \tfrac{1}{2}e^{(\lambda+\rho)}.$$

To obtain the lower bound in (6), consider the segment of $u_n = \lambda + \rho$ taken between $v_n = 0$ and $v_n = |n|$. According to (2), the H-length of this segment is $|n| \cosh (\lambda + \rho)$. Now

$$\rho = |AB| < |n| \cosh (\lambda + \rho),$$

since AB is the H-perpendicular let fall from A to $v_n = |n|$. The lower bound given in (6) is then obtained from this inequality, inasmuch as $|n| \cosh (\lambda + \rho) < |n| e^{(\lambda+\rho)}$.

LEMMA 6. *When the H-line segment AB lies in the region*

$$T_n: \qquad 0 < u_n \leqq \tau, \qquad 0 \leqq v_n \leqq |n|$$

of Φ and A lies on $u_n = 0$, we have

$$|AB| < |n| + \tau.$$

If AB lies in the region $-\mu \leqq u_n \leqq \tau$, $0 \leqq v_n \leqq |n|$, with A on $u_n = -\mu$ and B in T_n, we have

$$|AB| < |n| + \mu + \tau.$$

For the first part of the lemma, take C on $u_n = 0$, so that its v_n-coördinate equals that of B. Now

$$|AB| \leqq |AC| + |CB| \leqq |n| + \tau,$$

which proves this part of the lemma, inasmuch as the equality signs cannot hold simultaneously.

To prove the second part of the lemma, we take C as before and again have $|AB| \leqq |AC| + |CB|$, and since $|AC| < |n| + \mu$, $|CB| \leqq \tau$, the second part of the lemma is proved.

LEMMA 7. *When* $\sinh(|n|/2) < \tanh \delta$, *an H-circle of H-radius δ can be inscribed in the region T_n of Lemma 6 if $\tau \geqq \tau_n$, where*

$$\tau_n = \delta + \log \frac{4 \sinh \delta}{|n|}. \tag{7}$$

Let A be a point on the positive half of $v_n = |n|/2$, B be the foot of the H-perpendicular let fall from A to $v_n = |n|$, C be the point $[0, |n|]$, and D be the point $[0, |n|/2]$. An H-circle of H-radius $|AB|$ described about A as an H-center is tangent to both[8] $v_n = 0$, $v_n = |n|$ and will, in addition, be tangent to $u_n = 0$ if $|AD| = |AB|$. If, however, $|AD| > |AB|$, the H-circle lies in the region

$$0 < u_n \leqq |AB| + |AD|, \qquad 0 \leqq v_n \leqq |n|, \tag{8}$$

of Φ.

From (4) we have

$$\sinh |AB| = \sinh(|n|/2) \cosh |AD| < \tanh \delta \cosh |AD|,$$

and therefore the condition that $|AD| > |AB|$ is met, when $|AB| = \delta$. By way of obtaining an upper bound for $|AD|$, we have

$$\cosh |AD| = \frac{\sinh \delta}{\sinh(|n|/2)},$$

and therefore

$$|AD| = \log \left\{ \frac{\sinh \delta + [\sinh^2 \delta - \sinh^2(|n|/2)]^{\frac{1}{2}}}{\sinh(|n|/2)} \right\},$$

so that

$$|AD| < \log \frac{2 \sinh \delta}{\sinh(|n|/2)} < \log \frac{4 \sinh \delta}{|n|}.$$

[8] To show that it is tangent to $v_n = 0$, let B' be the foot of the H-perpendicular let fall from A to $v_n = 0$, C' be the point $[0, 0]$. From (4) we have, since $|C'D| = |CD|$, the result $\sinh |AB'| = \sinh |C'D| \cdot \cosh |AD| = \sinh |CD| \cosh |AD| = \sinh |AB|$. Hence $|AB'| = |AB|$, which completes the proof.

The upper bound in the first inequality in (8) may then be replaced by $\delta + \log \dfrac{4 \sinh \delta}{|n|}$. This proves the property stated for τ_n in the lemma.

LEMMA 8. *The H-length of the mutual H-perpendicular n between two non-intersecting H-lines l_1, l_2 drawn tangent to an H-circle Γ of H-radius γ is given by*

$$(9) \qquad | n | = 2 \log \{\sin \theta \cosh \gamma + (\sin^2\theta \cosh^2\gamma - 1)^{\frac{1}{2}}\},$$

where 2θ [arc sin (sech γ) $< \theta \leqq \pi/2$] is the angle between the H-radii of Γ drawn to the points of tangency.

Let A be the H-center of Γ, B be the point of tangency of Γ with l_1, C the end-point of n on l_1, and D the H-mid-point of n. From (5) we have

$$\cosh (| n |/2) = \sin \theta \cosh \gamma,$$

since $| CD | = | n |/2$, $\angle BAD = \theta$, $| AB | = \gamma$, and (9) follows from this equation.

When $0 \leqq \theta \leqq$ arc sin (sech γ), the two H-lines l_1, l_2 intersect.

LEMMA 9. *Let a_1, a_2, b_1, b_2 be four H-lines each making an angle θ $(-\pi/2 < \theta < \pi/2)$ with an H-line segment AB. Let those H-lines labeled with a intersect at A, and those labeled with b intersect at B. Choose θ so that no line labeled with a intersects a line labeled with b. The four H-distances*

$$| a_1b_1 |, \qquad | a_1b_2 |, \qquad | a_2b_1 |, \qquad | a_2b_2 |$$

resolve into two pairs such that the H-distances in each pair are equal, the greater H-distances occurring for the pair in which the H-lines are placed so that their mutual H-perpendicular intersects AB.

The H-distances from the H-mid-point O of AB to a_1, a_2, b_1, b_2, are all equal to one another, so that these four H-lines are all tangent to the same H-circle with H-center at O. The proof of this lemma is then readily seen to follow from (9) in Lemma 8.

4. **The domain D_0.** Let p denote a positive integer greater than unity and let ρ, δ $(0 < \rho < \delta)$ be defined by[9]

$$(10) \qquad \cosh \frac{\rho}{2} = \cot \frac{\pi}{4p},$$

$$(11) \qquad \cosh \frac{\delta}{2} = \cot^2 \frac{\pi}{4p}.$$

In Φ construct the H-circles

$$(12) \qquad x^2 + y^2 = \tanh^2 \frac{\rho}{4},$$

$$(13) \qquad x^2 + y^2 = \tanh^2 \frac{\delta}{4},$$

[9] I am indebted to the referee for the elegant form of these equations.

having H-radii $\rho/2$ and $\delta/2$ respectively, and beginning with the point on (12) where it intersects the positive x-axis, divide the circumference of (12) into $4p$ equal arcs. At each of the $4p$ division points draw an H-line tangent to (12), thereby forming a curvilinear polygon[10] having $4p$ *sides* and $4p$ *vertices*. We denote the interior of this curvilinear polygon by D_0 and label its sides

$$e_1, e_2, \cdots, e_{4p}, \tag{14}$$

e_1 being taken as that side of D_0 which is tangent to (12) where it cuts the positive x-axis, and the subscripts increasing when the boundary of D_0 is traversed counter-clockwise. The vertices of D_0 are labeled

$$V_1, V_2, \cdots, V_{4p}, \tag{15}$$

V_1 being the vertex at which e_1 and e_2 concur, and the subscripts increasing when the boundary of D_0 is traversed counter-clockwise. The vertices of D_0 all lie on (13). The H-circle (12) is inscribed in D_0, which in turn is inscribed in (13). The H-length of each side of D_0 equals ρ and the interior angle at each vertex equals $\pi/2p$.

V_j and V_{j-1} ($V_0 = V_{4p}$) each divide the side e_j of D_0 produced into two H-rays. The H-ray ending on V_j (V_{j-1}) which does not contain the side e_j of D_0 is denoted by e_j^+ (e_j^-). It is assumed that V_j (V_{j-1}) is a point of e_j^+ (e_j^-). The region of Φ lying outside of (13) is divided into two subregions by e_{j+1}^-, e_{j-1}^+ ($e_{4p+1} = e_1$, $e_0 = e_{4p}$). The subregion bounded on Ψ by the shorter (Euclidean) arc when taken with its boundary is denoted by E_j. The region of Φ, lying outside of (13) and between E_j and E_{j+1} ($E_{4p+1} = E_1$) at V_j, when taken with its boundary is denoted by E_j'. With respect to the region E_j, we prove the following lemma.

Lemma 10. *The boundaries e_{j+1}^-, e_{j-1}^+ of E_j cannot be joined by an H-line segment lying in E_j.*

The proof of this lemma is elementary, but because of the importance of the lemma in subsequent arguments the proof is given in some detail. From symmetry considerations it is clear that we need consider only the case $j = 1$.

In the x, y-coördinate system the equation of the circle containing the side e_2 of D_0 as a segment is

$$x^2 + y^2 - 2\sqrt{\cos\frac{\pi}{2p}}\cos\frac{\pi}{4p}x - 2\cos\frac{\pi}{4p}\sqrt{\sec\frac{\pi}{2p}}\sin\frac{\pi}{2p}y + 1 = 0.$$

The minimum value of y on this circle is calculated to be $\sin \pi/(4p)\sqrt{\cos \pi/(2p)}$, occurring, as a matter of fact, at V_1. Consider the H-line l drawn tangent to (13) at the point where it cuts the positive x-axis. l is a segment of the circle whose equation is

$$x^2 + y^2 - \sqrt{\sec\frac{\pi}{2p}}\left(1 + \cos\frac{\pi}{2p}\right)x + 1 = 0,$$

[10] That the H-lines so constructed intersect in pairs is well known. This fact may also be drawn from (10) in connection with the remark at the end of Lemma 8 [Referee].

and lies exterior to (13). The maximum value of y on l is $\tan^2\pi/(4p)$ and it is readily seen that

$$\tan^2 \frac{\pi}{4p} < \sin \frac{\pi}{4p} \sqrt{\cos \frac{\pi}{2p}} \qquad (p = 2, 3, \cdots),$$

so that l cannot intersect the side e_2 of D_0 produced, the situation being analogous in regard to the side e_{4p} of D_0. l therefore lies in E_1.

Suppose an H-line segment lies in E_1 and connects e_2^-, e_{4p}^+. Since l connects (13) to Ψ, this H-line segment meets l. It cannot, however, be a segment of l and therefore when produced intersects Ψ in two points, one in each of the arcs into which Ψ is divided by the end-points of l, so that it cannot intersect both e_2^-, e_{4p}^+, which is a contradiction. The lemma is therefore true.

5. **The group G and the net N.** When certain of the points on the boundary of D_0 are adjoined to D_0, a fundamental domain $\bar{D}_0$ is obtained for a well-known[11] Fuchsian group G. Transformations of G which take the sides of D_0 into one another in the following manner:

$$e_1 \rightarrow e_3,\ e_4 \rightarrow e_2;\ e_5 \rightarrow e_7,\ e_8 \rightarrow e_6;\ \cdots;\ e_{4p-3} \rightarrow e_{4p-1},\ e_{4p} \rightarrow e_{4p-2},$$

form a set of generators for G. The transformations of G are all H-transformations.

Two point sets in Φ, or two sets of directed H-lines, or two sets of directed H-line segments which can be transformed into one another by transformations of G are *congruent*, or *copies* of one another; two symbols forming a pair above are *congruent symbols*. The copies of $\bar{D}_0$ are *meshes* and their totality covers Φ without lacunae. The copies of the sides (vertices) of D_0 are the *sides* (*vertices*) *of the meshes*. Two meshes with a side in common *border* upon the common side. Those meshes which border $\bar{D}_0$ upon $e_1, e_2, \cdots, e_{4p}$ are denoted by $\bar{D}_1, \bar{D}_2, \cdots, \bar{D}_{4p}$ respectively. Two vertices which are end-points of the same side of a mesh are *adjacent vertices*. Two sides of a mesh concurring at a vertex of the mesh are *adjacent sides*. Two sides of a mesh which are not adjacent sides are *non-adjacent sides*. Two non-adjacent sides of a mesh adjacent to the same side of the mesh are *alternating sides*. Alternating sides, as well as adjacent sides, always belong to the same mesh. When two meshes border upon a common side, this common side is adjacent to two pairs of alternating sides, one pair being found in each mesh. If a side be selected from each of these pairs of alternating sides so that the two sides selected do not concur, the two sides thus selected are *opposite sides*. Opposite sides do not belong to the same mesh but belong to meshes which have a side in common. Neither non-adjacent sides nor opposite sides intersect, even when produced.

The H-length of each side of a mesh equals ρ. Any given mesh is inscribed

[11] See, for example, M. Morse, Trans. Amer. Math. Soc., vol. 26 (1924), pp. 25–32. This paper will be referred to as Morse II. See also, J. Nielsen, Acta. Math., vol. 50 (1927), pp. 191–224.

in an H-circle of H-radius $\delta/2$ and its $4p$ sides are all tangent to an H-circle of H-radius $\rho/2$. The interior angle at each vertex of the mesh equals $\pi/2p$.

The sides of the meshes form a *network* N of H-line segments. The end-points of the sides in N are the *vertices* of N. The sides in N align themselves into H-lines, $2p$ of these H-lines concurring at each vertex of N. A vertex of N therefore serves as a vertex for $4p$ meshes. We label each side of a mesh with the label assigned in (14) to that side of D_0 into which it is carried by the transformation of G carrying the mesh into $\bar{D}_0$. To each side of a mesh there is then assigned two of the symbols in (14), one coming from each of the meshes bordering upon the side in question. N, together with the labeling of its sides, is transformed into itself by a transformation of G.

The two points in which the side e_i of D_0 produced intersects Ψ are the *base-points* of e_i. The base-points of e_i divide Ψ into two arcs of unequal euclidean length and e_i *subtends* the shorter arc (end-points included). Let $[e_i]$ denote the arc on Ψ which e_i subtends. Adjacent sides of D_0 subtend overlapping arcs on Ψ and the part of $[e_i]$ remaining when the overlapping parts are removed is denoted by $[e_i]_1$. Obviously $[e_i]_1$ and $[e_j]_1$ have no common point, unless $i = j$. The transforms of the arcs $[e_i]$, $[e_i]_1$ ($i = 1, 2, \cdots, 4p$) by the transformation which carries $\bar{D}_0$ into one of $\bar{D}_1, \bar{D}_2, \cdots, \bar{D}_{4p}$, say $\bar{D}_k$, are denoted by $[e_i]^{(k)}$, $[e_i]_1^{(k)}$ respectively. If we follow the conventions used above, the side e_i of $\bar{D}_k$ subtends the arc $[e_i]^{(k)}$ on Ψ and the end-points of $[e_i]^{(k)}$ are the base-points of the side e_i of $\bar{D}_k$.

We now prove four lemmas.

LEMMA 11. *The H-length of an H-line segment s whose end-points lie on two non-adjacent sides, or on two opposite sides, cannot fall below the constant χ, where*

$$\chi = 2 \log\left\{2 \cos^2 \frac{\pi}{4p} + \left(4 \cos^4 \frac{\pi}{4p} - 1\right)^{\frac{1}{2}}\right\}, \tag{16}$$

and is actually greater than χ in the latter case.

First, suppose the end-points of s rest upon two non-adjacent sides. $|s|$ cannot fall below the H-distance between the two non-adjacent sides produced.

If we identify these two non-adjacent sides produced as l_1, l_2 in Lemma 8 and the H-circle of H-radius $\rho/2$ to which they are drawn tangent as Γ, the H-distance between two non-adjacent sides produced is readily seen to be least when the two non-adjacent sides are alternating sides, inasmuch as $|n|$ in (9) is a monotone increasing function of θ in the given interval. Let χ denote the H-distance between two alternating sides produced. We have $|s| \geqq \chi$, and in order to obtain the value of χ given in (16), we place $\theta = \pi/2p$, $\gamma = \rho/2$ in (9), ρ being taken as given in (10).

Second, suppose the end-points of s rest upon two opposite sides. According to the definition of opposite sides given above, the end-points of s lie upon two sides of N which do not belong to the same mesh but to two meshes bordering on a side, say AB of N. One of the opposite sides emanates from A and the other from B, the angles measured from AB to the sides being simultaneously

$\pi/2p$ or $-\pi/2p$. The two opposite sides produced may now be interpreted as a pair selected from the four H-lines a_1, a_2, b_1, b_2 in Lemma 9; one of the opposite sides produced will be labeled with a, the other with b. From Lemma 9 it follows that the distance between two opposite sides produced exceeds the distance between two alternate sides produced. Hence when the end-points of s rest on two opposite sides, $|s|$ exceeds χ.

LEMMA 12. *If k denotes any fixed positive integer taken from $1, 2, \cdots, 4p$, exactly $4p - 3$ sides of the mesh $\bar{D}_k$ have both their base-points in the interior of $[e_k]_1$. If these $4p - 3$ sides are taken in their circular order around $\bar{D}_k$, consecutive sides are adjacent, except two, between which the three remaining sides of $\bar{D}_k$ intervene.*

For simplicity we shall prove this lemma for $p = 2$, the proof permitting an obvious extension to any value of p. In addition we may restrict ourselves with no loss in generality to the case $k = 1$. The meshes $\bar{D}_0$ and $\bar{D}_1$ border along a side of N which when reckoned to $\bar{D}_0$ ($\bar{D}_1$) is labeled e_1 (e_3). If we proceed from this side in a counter-clockwise fashion around the periphery of $\bar{D}_1$, the labels of the sides of $\bar{D}_1$ when reckoned to $\bar{D}_1$ are met in order $e_3, e_4, e_5, e_6, e_7, e_8, e_1, e_2, e_3$.

We now prove that the sides e_5, e_6, e_7, e_8, e_1 of $\bar{D}_1$ have both base-points in the interior of $[e_1]_1$.

Let us begin with e_5. Suppose both base-points of e_5 do not lie in $[e_1]_1$. Since the sides e_5, e_3 of $\bar{D}_1$ when produced cannot intersect, the only way this can occur is for the side e_5 of $\bar{D}_1$ produced to meet the side e_8 of $\bar{D}_0$ produced. On interpreting the H-line obtained by producing the side e_4 of $\bar{D}_1$ as a transversal which cuts across the two H-lines obtained by producing the side e_8 of $\bar{D}_0$ and the side e_5 of $\bar{D}_1$, we easily see from elementary non-Euclidean geometry that the latter two H-lines cannot intersect. This follows from the fact that the sum of the interior angles lying exterior to $\bar{D}_1$ is for any p equal to $\pi(2 - 3/2p) \geqq \pi(2 - 3/4) > \pi$.

In like manner it may be shown that the base-points of the side e_1 of $\bar{D}_1$ lie in $[e_1]_1$.

The truth of the statement for the remaining sides e_6, e_7, e_8 of $\bar{D}_1$ is now obvious.

Let Ω denote an arc on Ψ which is the sum of $4p - 4$ consecutive $[e_j]$'s and let $\mathrm{A} = \Psi - \Omega$. When $\bar{D}_0$ is transformed into a given $\bar{D}_i$ ($i = 1, 2, \cdots, 4p$) by a transformation of G, the transform of A under this transformation will be denoted by $\mathrm{A}^{(i)}$.

LEMMA 13. *There are at least two different values of i such that $[e_i]_1 \supset \mathrm{A}^{(i)}$.*

In view of the symmetry of D_0, it is sufficient to prove the lemma when

$$\mathrm{A} = [e_{4p-3}]_1 + [e_{4p-2}] + [e_{4p-1}] + [e_{4p}]_1, \qquad \Omega = \sum_{\nu=1}^{4p-4} [e_\nu].$$

Here $\mathrm{A} \subset [e_{4p-3}] + [e_{4p-2}] + [e_{4p-1}] + [e_{4p}]$, so that $[e_i]_1 \supset \mathrm{A}^{(i)}$ whenever each of the sides e_{4p-3}, e_{4p-2}, e_{4p-1}, e_{4p} of $\bar{D}_i$ has both base-points in $[e_i]_1$.

According to Lemma 12, $4p - 3$ consecutive sides of $\bar{D}_i$ have both base-points in $[e_i]_1$. From these $4p - 3$ sides $4p - 6$ combinations of four consecutive sides each can be formed. In all the meshes $\bar{D}_1, \bar{D}_2, \cdots, \bar{D}_{4p}$ there are $4p(4p - 6)$ such combinations, not all different to be sure, but any one combination occurring as often as any other. Since $4p$ different combinations of four consecutive sides are possible, *any one combination* such as $e_{4p-3}, e_{4p-2}, e_{4p-1}, e_{4p}$ occurs $4p(4p - 6)/4p = 4p - 6$ times among the $4p(4p - 6)$ combinations. Now a given combination of four consecutive sides occurs only once in a given $\bar{D}_i$, so that the combination $e_{4p-3}, e_{4p-2}, e_{4p-1}, e_{4p}$ occurs in the desired manner in $4p - 6$ different $\bar{D}_i$'s. The number of values of i such that $[e_i]_1 \supset A^{(i)}$ is at least $4p - 6 \geqq 2$, since $p \geqq 2$.

LEMMA 14. *When* $\sinh(|n|/2) < \tanh \delta$ *and if* $\tau \geqq \tau_n$, *where* τ_n *is given in* (7), *the region* T_n *in Lemma* 6 *contains at least one entire mesh, together with the H-circle in which it is inscribed.*

A region of Φ comprising the interior and boundary of an H-circle of H-radius δ drawn about an arbitrary point of Φ as H-center contains at least one entire mesh together with the H-circle in which it is inscribed.[12] In view of this fact, and Lemma 7, this lemma is obvious.

6. **Admissible sequences.** Let A, B be two points of Φ, not lying on N, such that the H-line segment AB connecting them intersects k sides of N without passing through a vertex of N. The directed H-line segment $\overrightarrow{AB}$ generates a sequence of the e's in the following way: beginning at A, we proceed along $\overrightarrow{AB}$ until we meet a side in N labeled a_1, say, when the side is reckoned to the mesh we are just leaving; after leaving this mesh, we continue along $\overrightarrow{AB}$ until we meet a second side of N labeled, say, a_2, when the side is reckoned to the mesh we are just leaving; etc. The process ends when B is reached and yields a certain sequence

$$a_1 a_2 \cdots a_{k-1} a_k \tag{17}$$

of the e's which is the *admissible sequence of* $\overrightarrow{AB}$. The admissible sequence of the oppositely directed H-line segment $\overrightarrow{BA}$ is then

$$a_k' a_{k-1}' \cdots a_2' a_1',$$

where $a_k', a_{k-1}', \cdots, a_2', a_1'$ designate symbols congruent to $a_k, a_{k-1}, \cdots, a_2, a_1$ respectively.

If $\overrightarrow{AB}$ is a segment of $\overrightarrow{A_1B_1}$ and if they meet N at the same points, their admissible sequences are identical. Here neither $\overrightarrow{AB}$ nor $\overrightarrow{A_1B_1}$ is supposed to meet a vertex of N or to end upon N. The *admissible sequence of* $\overrightarrow{AB}$ *where A and B are points of N but not vertices of N* will be defined to be the admissible sequence of $\overrightarrow{A_1B_1}$, where A_1, B_1 are two points not lying on N, such that $\overrightarrow{A_1B_1}$

[12] Let P denote an arbitrary point of Φ. P necessarily lies in at least one mesh, which in turn is inscribed in an H-circle Γ of H-radius $\delta/2$. An H-circle of H-radius δ taken about P as an H-center will contain Γ.

contains $\overrightarrow{AB}$ as a segment, and meets N only at its intersection points with $\overrightarrow{AB}$.

For brevity we denote the side in N which gives rise to a_i ($i = 1, 2, \cdots, k$) in (17) by a_i. The segment of $\overrightarrow{AB}$ comprehended between two sides a_{i_1}, a_{i_2} ($i_1 < i_2$) of N is conveniently denoted by $[a_{i_1} a_{i_2}]$.

We now prove three lemmas related to admissible sequences.

LEMMA 15. *If A, B are two points on N, not vertices of N, such that $\overrightarrow{AB}$ does not pass through a vertex of N, and if the admissible sequence of $\overrightarrow{AB}$ contains $2pr + 1$ symbols, where r is a positive integer, $|AB|$ is bounded from below by*

$$r\chi \leqq |AB|, \tag{18}$$

where χ is defined in (16).

Let

$$a_1 a_2 \cdots a_{2p} a_{2p+1} \cdots a_{2p\nu+1} a_{2p\nu+2} \cdots a_{2p(\nu+1)} a_{2p(\nu+1)+1} \cdots a_{2pr+1} \tag{19}$$

be the admissible sequence of $\overrightarrow{AB}$ and pick out from it the subsequences

$$a_{2p\nu+1} a_{2p\nu+2} \cdots a_{2p(\nu+1)} a_{2p(\nu+1)+1} \qquad (\nu = 0, 1, \cdots, r-1)$$

each containing $2p + 1$ symbols. The proof of the lemma is achieved by showing that none of the H-lengths

$$|[a_{2p\nu+1} a_{2p(\nu+1)+1}]| \qquad (\nu = 0, 1, \cdots, r-1),$$

can fall below χ. Let us prove this for $\nu = 0$, the method of proof being the same for the remaining values of ν. Consider the subsequence $a_1 a_2 \cdots a_{2p} a_{2p+1}$. The end-points of $[a_1 a_{2p+1}]$ lie on N but are not vertices of N.

If at least one of the $2p$ segments into which $[a_1 a_{2p+1}]$ is divided by the $2p$ meshes through which it passes has its end-points on non-adjacent sides, we have $|[a_1 a_{2p+1}]| \geqq \chi$ by virtue of Lemma 11.

If, however, each of these $2p$ segments has its end-points on adjacent sides, we proceed as follows: We first note that two sides in N from which two consecutive symbols in $a_1 a_2 \cdots a_{2p} a_{2p+1}$ arise meet at a vertex of N. Let $V^{(1)}, V^{(2)}, \cdots, V^{(2p)}$ denote the vertices (some of which may coincide) where the sides a_1, a_2; a_2, a_3; $\cdots$; a_{2p}, a_{2p+1} of N respectively concur. At least two of these vertices do not coincide. For suppose they all coincide. The $2p$ meshes through which $[a_1 a_{2p+1}]$ passes then possess $V^{(1)}$ as a vertex and their interior angles at $V^{(1)}$ would exhaust a straight angle at $V^{(1)}$. $[a_1 a_{2p+1}]$ intersects both sides of this straight angle, and therefore intersects an H-line passing through $V^{(1)}$ in two distinct points, thus coinciding with a segment of this H-line. This requires that $[a_1 a_{2p+1}]$ pass through a vertex of N, contrary to our assumption for $\overrightarrow{AB}$. Hence all the vertices cannot coincide. Let $V^{(j)}$ $(1 < j \leqq 2p)$ be the first vertex which does not coincide with $V^{(1)}$. $V^{(j)}$ and $V^{(1)}$ are adjacent vertices and are the end-points of the side a_j of N. Now consider the H-line segment $[a_{j-1} a_{j+1}]$. The end-points of this H-line segment lie on opposite sides[13] and

[13] For the definition of opposite sides see §5, or the proof of the second part of Lemma 11.

therefore from Lemma 11, we have $|[a_{j-1}a_{j+1}]| > \chi$, from which it follows *a fortiori* that $|[a_1 a_{2p+1}]| > \chi$.

LEMMA 16. *If $\overrightarrow{A_0B_0}$, $\overrightarrow{A_1B_1}$ are two directed H-line segments such that A_0, A_1 lie on the same side in N, and B_0, B_1 lie on the same side in N, and if neither of $\overrightarrow{A_0B_0}$, $\overrightarrow{A_1B_1}$ meets a vertex of N, their admissible sequences contain the same number of symbols.*

Let a (b) denote the side in N on which rest A_0, A_1 (B_0, B_1) and let α (β) denote the H-distance along a (b) measured from one of the end-points [arbitrarily fixed] of a (b). A directed H-line segment with α-end-point on a and ω-end-point on b determines a pair α, β uniquely, and vice versa. Let α_0, β_0 be the pair determined by $\overrightarrow{A_0B_0}$ and α_1, β_1 the pair determined by $\overrightarrow{A_1B_1}$. The directed H-line segment $\overrightarrow{A_tB_t}$ determined by the pair α_t, β_t, where

$$\alpha_t = \alpha_0 + t(\alpha_1 - \alpha_0), \qquad \beta_t = \beta_0 + t(\beta_1 - \beta_0) \qquad (0 \leqq t \leqq 1),$$

is deformed from $\overrightarrow{A_0B_0}$ into $\overrightarrow{A_1B_1}$ when t ranges from 0 to 1, and its *end-points* A_t, B_t are displaced along a, b respectively, without meeting a vertex of N.

Suppose $\overrightarrow{A_0B_0}$ intersects N in k points. None of these points is a vertex of N and therefore $\overrightarrow{A_0B_0}$ intersects k H-lines in N. Let P denote an intersection point of $\overrightarrow{A_0B_0}$ with one of these k H-lines, say with the H-line l. During the deformation of $\overrightarrow{A_tB_t}$ from $\overrightarrow{A_0B_0}$ to $\overrightarrow{A_1B_1}$ the point P moves along both $\overrightarrow{A_tB_t}$ and l. At no time can P pass off the end-points of l, since they rest on Ψ. Likewise it cannot pass off the end-points of $\overrightarrow{A_tB_t}$, since this would require that either A_t or B_t coincide with a vertex of N; which is contrary to hypothesis. Conceivably the point P might be lost by $\overrightarrow{A_tB_t}$ and l becoming tangent to each other. This, however, cannot occur for it would imply that $\overrightarrow{A_tB_t}$ is contained in l and therefore that A_t, B_t are vertices of N. At the conclusion of the deformation, $\overrightarrow{A_1B_1}$ therefore intersects N in at least k points, none of which can coincide, since $\overrightarrow{A_1B_1}$ does not meet a vertex of N. On interchanging the rôles of $\overrightarrow{A_0B_0}$ and $\overrightarrow{A_1B_1}$ in the deformation, we see that $\overrightarrow{A_0B_0}$ and $\overrightarrow{A_1B_1}$ intersect N in the same number of points. The number of symbols in each admissible sequence is therefore equal to k.

LEMMA 17. *The H-length of an H-line segment $\overrightarrow{A_1B_1}$ which terminates on the same sides of N as does the H-line segment $\overrightarrow{AB}$ in Lemma* 15 *cannot fall below* $r\chi$.

If $\overrightarrow{A_1B_1}$ does not pass through a vertex of N, according to Lemma 16 the number of symbols in its admissible sequence is the same as the number of symbols in the admissible sequence of $\overrightarrow{AB}$. Hence, from Lemma 15, $|A_1B_1| \geqq r\chi$.

This result holds, even if $\overrightarrow{A_1B_1}$ passes through a vertex of N. Note that there is a one-to-one correspondence between the points of the square $0 \leqq \alpha \leqq \rho$, $0 \leqq \beta \leqq \rho$ in the α, β-plane (where α, β are taken as introduced in the proof of Lemma 16) and the directed H-line segments $\overrightarrow{A_1B_1}$ which have an end-point upon each of the above sides. The points in this square which correspond to H-line segments passing through a vertex of N are limit points of the set of points in the square corresponding to H-line segments which do not pass through

a vertex of N. Now the H-length of $\overrightarrow{A_1B_1}$ is a continuous function of the arguments α_1, β_1, so that the H-length of $\overrightarrow{A_1B_1}$ in any position in which it passes through a vertex of N cannot fall below $r\chi$.

7. **Orthogonal families of directed H-lines.** Let n' be an arbitrarily given directed H-line segment. Following §2 we adopt the coördinate system $[u_{n'}, v_{n'}]$ on Φ associated with n'. The set of H-lines $v_{n'} = c$ $(0 \leqq c \leqq |n'|)$ directed from $u_{n'} = -\infty$ to $u_{n'} = +\infty$ is an *orthogonal family of directed H-lines* and is denoted by $F_{n'}$, the orthogonal family of oppositely directed H-lines being denoted by $\bar{F}_{n'}$. $F_{n'}$, $\bar{F}_{n'}$ are *based* on n'. The arc on Ψ occupied by the α (ω)-end-points of the directed H-lines in $F_{n'}$ is the *α-arc of $F_{n'}$* (*ω-arc of $F_{n'}$*) and is denoted by $\alpha_{n'}$ $(\omega_{n'})$. Let $a_{n'}$, $b_{n'}$ denote the two directed H-lines of $F_{n'}$ drawn through the end-points of n', choosing the notation so that when $\omega_{n'}$ is traversed counter-clockwise it leads from $a_{n'}$ to $b_{n'}$. According to Lemma 14, the region $T_{n'}$: $0 < u_{n'} \leqq \tau_{n'}$, $0 \leqq v_{n'} \leqq |n'|$, of Φ, where $\sinh (|n'|/2) < \tanh \delta$, $\tau_{n'} = \delta + \log 4 \sinh \delta/|n'|$, contains at least one entire mesh together with the H-circle in which it is inscribed, and may therefore be transformed by a transformation of G into a region of Φ which contains (13). The transforms of n', $F_{n'}$, $a_{n'}$, $\omega_{n'}$, $a_{n'}$, $b_{n'}$, $T_{n'}$ by this transformation are indicated by dropping the primes. Note that although either, or both, of a_n, b_n can be tangent to (13), neither intersects the interior of (13), and that since n' lies without $T_{n'}$, n lies without (13).

The transforms of n, F_n, α_n, ω_n, a_n, b_n, T_n by the transformation carrying $\bar{D}_0$ into a stated $\bar{D}_i$ $(i = 1, 2, \cdots, 4p)$ are indicated by adding the superscript (i) to n; thus $T_{n^{(i)}}$ is the transform of T_n. Note that we have

$$T_{n^{(i)}} \supset \bar{D}_i \qquad (i = 1, 2, \cdots, 4p). \tag{20}$$

We now prove three lemmas.

LEMMA 18. *ω_n contains at least $4p - 4$ consecutive $[e_j]$'s in its interior.*

For reasons of symmetry, it is sufficient to prove the lemma in the case when the ω-end-point of a_n lies in $[e_{4p}]$, but not in $[e_1]$; and we restrict our attention to this case.

From Lemma 10 it is seen that a_n lies in the region $E_{4p-1} + E'_{4p-1} + E_{4p}$. Suppose the lemma were false. The ω-end-point of b_n would have to lie in one of the arcs $[e_{4p-4}]$, $[e_{4p-5}]$, $\cdots$, $[e_1]$. Let us assume that the ω-end-point of b_n lies in $[e_{4p-4}]$. From Lemma 10 it follows that b_n lies in the region $E'_{4p-5} + E_{4p-4} + E'_{4p-4} + E_{4p-3}$. The H-line segment n which joins a_n and b_n lies outside of (13) and is therefore required to intersect the side e_{4p-2} of D_0 produced in two distinct points. This would necessitate that n contain the side e_{4p-2} of D_0 as a segment, which is impossible, since n does not intersect (13). In like manner, a contradiction may be reached in assuming the ω-end-point of b_n to lie in any one of $[e_{4p-5}]$, $[e_{4p-6}]$, $\cdots$, $[e_1]$. The lemma is therefore true.

LEMMA 19. *There are at least two different values of i such that $[e_i]_1$ contains $\alpha_{n^{(i)}}$ together with the end-points of $\omega_{n^{(i)}}$.*

According to Lemma 18, ω_n contains an arc Ω in its interior, α_n and the end-points of ω_n being contained in $\Psi - \Omega = \Lambda$. Hence $\alpha_{n^{(i)}}$ and the end-points of $\omega_{n^{(i)}}$ lie in $\Lambda^{(i)}$, which in turn, from Lemma 13, is contained in $[e_i]_1$, for at least two different values of i.

Let $F_{n_1'}$ denote an orthogonal family of directed H-lines which may, or may not, be identical with $F_{n'}$. We suppose $\sinh(|n_1'|/2) < \tanh\delta$, so that the region

$$\bar{T}_{n_1'}: \quad -\tau_{n_1'} \leqq u_{n_1'} < 0, \qquad 0 \leqq v_{n_1'} \leqq |n_1'|,$$

where $\tau_{n_1'}$ is defined by (7), contains at least one entire mesh together with the H-circle in which it is inscribed. We introduce

$$n_1^{(i)}, F_{n_1^{(i)}}, \alpha_{n_1^{(i)}}, \omega_{n_1^{(i)}}, a_{n_1^{(i)}}, b_{n_1^{(i)}}, \bar{T}_{n_1^{(i)}} \qquad (i = 1, 2, \cdots, 4p),$$

which are defined like $n^{(i)}, F_{n^{(i)}}, a_{n^{(i)}}, \omega_{n^{(i)}}, a_{n^{(i)}}, b_{n^{(i)}}, T_{n^{(i)}}$, the rôle of the region $T_{n'}$ above being taken now by $T_{n_1'}$, so that in place of (20) we have

$$\bar{T}_{n_1^{(i)}} \supset \bar{D}_i \qquad (i = 1, 2, \cdots, 4p). \tag{21}$$

The following companion lemma to Lemma 19 is then evident if we apply Lemma 19 to $\bar{F}_{n_1'}$.

Lemma 20. *There are at least two different values of i such that $[e_i]_1$ contains $\omega_{n_1^{(i)}}$, together with the end-points of $\alpha_{n_1^{(i)}}$.*

8. **δ-sets of directed H-line segments.** Let n' be an arbitrarily given directed H-line segment and let $[u_{n'}, v_{n'}]$ be the coördinate system on Φ used in §7 to define $F_{n'}$. A *δ-set of directed H-line segments* $\Delta_{n'}$, or briefly, a *δ-set* $\Delta_{n'}$, is defined to be the totality of directed H-line segments having their α (ω)-end-points on the segment $0 \leqq v_{n'} \leqq |n'|$ of $u_{n'} = -\delta$ ($u_{n'} = \delta$). $\Delta_{n'}$ is *based* on n', and $\Delta_{n'}$, $F_{n'}$ are said to be *associated.*

We now prove the following lemma.

Lemma 21. *Let $\Delta_{n'}$ $[\Delta_{n_1'}]$ be a δ-set based on the directed H-line segment n' $[n_1']$, where* $\sinh(|n'|/2) < \tanh\delta$ $[\sinh(|n_1'|/2) < \tanh\delta]$. *Let $F_{n'}$ $[F_{n_1'}]$ be the orthogonal family of directed H-lines based on n' $[n_1']$ and associated with $\Delta_{n'}$ $[\Delta_{n_1'}]$. Let $\Delta_{n^{(i)}}$ $[\Delta_{n_1^{(i)}}]$ $(i = 1, 2, \cdots, 4p)$ denote the δ-set associated with the orthogonal family $F_{n^{(i)}}$ $[F_{n_1^{(i)}}]$ of directed H-lines obtained from $F_{n'}$ $[F_{n_1'}]$ in* §7, *noting that $\Delta_{n^{(i)}}$ $[\Delta_{n_1^{(i)}}]$ is a copy of $\Delta_{n'}$ $[\Delta_{n_1'}]$.*

Among the $F_{n^{(i)}}$ $(i = 1, 2, \cdots, 4p)$ there exists one, say $F_{n^{(k)}}$, which contains a set $S_{n^{(k)}}$ of directed H-line segments having the following properties:

I. *Each member of $S_{n^{(k)}}$ intersects*[14] *$\Delta_{n^{(k)}}$ and there exists a $\Delta_{n_1^{(i)}}$, say $\Delta_{n_1^{(k_1)}}$, which is intersected by all the members of $S_{n^{(k)}}$.*

II. *The α-end-points of the members of $S_{n^{(k)}}$ lie on the H-line segment*

$$u_{n^{(k)}} = -3\delta, \qquad 0 \leqq v_{n^{(k)}} \leqq |n'|, \tag{22}$$

[14] Cf. §1.

and the H-length of each member is less than λ_1, *where*

$$\lambda_1 = 11\delta + \log \frac{(4 \sinh \delta)^2}{|n'|\,|n_1'|}. \tag{23}$$

III. *The members of* $S_{n^{(k)}}$ *intersect* $n^{(k)}$ *in the points of an arc* μ_1, *which contains an arc* ν_1' *whose H-length is given by*

$$|\nu_1'| = |n_1'|\, e^{-\lambda_1}. \tag{24}$$

Property I. Following Lemma 19, choose k from $1, 2, \cdots, 4p$ so that $[e_k]_1$ contains $\alpha_{n^{(k)}}$ and the end-points of $\omega_{n^{(k)}}$. This may be done in at least two ways. Having chosen k, take k_1 so that $[e_{k_1}]_1$ contains $\omega_{n_1^{(k_1)}}$ and the end-points of $\alpha_{n_1^{(k_1)}}$. According to Lemma 20, this may be done in at least two ways, so that we can realize $k_1 \neq k$. Since $[e_k]_1$, $[e_{k_1}]_1$ $(k \neq k_1)$ have no common points, we see that

$$\omega_{n_1^{(k_1)}} \text{ and the end-points of } \alpha_{n_1^{(k_1)}} \subset \text{interior of } \omega_{n^{(k)}}, \tag{25}$$

$$\alpha_{n^{(k)}} \text{ and the end-points of } \omega_{n^{(k)}} \subset \text{interior of } \alpha_{n_1^{(k_1)}}. \tag{26}$$

Associated with $n^{(k)}$, $n_1^{(k_1)}$ there are the coördinate systems $[u_{n^{(k)}}, v_{n^{(k)}}]$, $[u_{n_1^{(k_1)}}, v_{n_1^{(k_1)}}]$ respectively. $n^{(k)}$ $(n_1^{(k_1)})$ is the segment $0 \leqq v_{n^{(k)}} \leqq |n'|$ $(0 \leqq v_{n_1^{(k_1)}} \leqq |n_1'|)$ of $u_{n^{(k)}} = 0$ $(u_{n_1^{(k_1)}} = 0)$ directed in the sense of increasing $v_{n^{(k)}}$ $(v_{n_1^{(k_1)}})$ and the directed H-lines in $F_{n^{(k)}}$ $(F_{n_1^{(k_1)}})$ are the coördinate lines $v_{n^{(k)}} = d$ $(v_{n_1^{(k_1)}} = d_1)$, where $0 \leqq d \leqq |n'|$ $(0 \leqq d_1 \leqq |n_1'|)$, directed from $u_{n^{(k)}} = -\infty$ $(u_{n_1^{(k_1)}} = -\infty)$ to $u_{n^{(k)}} = +\infty$ $(u_{n_1^{(k_1)}} = +\infty)$.

Using (25), we may show that $u_{n_1^{(k_1)}} = 0$ lies in the region

$$0 < u_{n^{(k)}} < +\infty, \qquad 0 < v_{n^{(k)}} < |n'|, \tag{27}$$

of Φ, and therefore that $u_{n_1^{(k_1)}} = c$ $(c > 0)$ lies in (27). Conversely, using (26), we find that $u_{n^{(k)}} = 0$ lies in the region

$$-\infty < u_{n_1^{(k_1)}} < 0, \qquad 0 < v_{n_1^{(k_1)}} < |n_1'| \tag{28}$$

of Φ, and therefore that $u_{n^{(k)}} = -c$ $(c > 0)$ lies in (28).

Consider the δ-sets $\Delta_{n^{(k)}}$, $\Delta_{n_1^{(k_1)}}$. Here $\Delta_{n^{(k)}}$ $(\Delta_{n_1^{(k_1)}})$ is the set of directed H-line segments having their α- and ω-end-points respectively on the segments of $u_{n^{(k)}} = -\delta$, $u_{n^{(k)}} = \delta$ $(u_{n_1^{(k_1)}} = -\delta$, $u_{n_1^{(k_1)}} = \delta)$ for which $0 \leqq v_{n^{(k)}} \leqq |n'|$ $(0 \leqq v_{n_1^{(k_1)}} \leqq |n_1'|)$. In order that a directed H-line l intersects $\Delta_{n_1^{(k_1)}}$, it is sufficient that the α-end-point of l lies on $\alpha_{n_1^{(k_1)}}$ and that l intersects the H-line segment

$$u_{n_1^{(k_1)}} = \delta, \qquad 0 \leqq v_{n_1^{(k_1)}} \leqq |n_1'|. \tag{29}$$

Any directed H-line in $F_{n^{(k)}}$ intersects $\Delta_{n^{(k)}}$, and since $\alpha_{n^{(k)}} \subset \alpha_{n_1^{(k_1)}}$, those directed H-lines in $F_{n^{(k)}}$ which intersect (29) will intersect $\Delta_{n_1^{(k_1)}}$. That directed H-lines in $F_{n^{(k)}}$ exist which intersect (29) is trivial, since $u_{n_1^{(k_1)}} = \delta$ lies in (27), and through any given point of (27) there passes exactly one directed H-line of $F_{n^{(k)}}$.

A set $S_{n^{(k)}}$ of directed H-line segments which has property I of the lemma can then be extracted from directed H-lines of $F_{n^{(k)}}$.

Property II. Let l denote an arbitrary member of $F_{n^{(k)}}$ which intersects $\Delta_{n_1^{(k_1)}}$ in addition to intersecting $\Delta_{n^{(k)}}$. Let A, B, C (A_1, B_1, C_1) denote the points where l intersects $u_{n^{(k)}} = -\delta$, $u_{n^{(k)}} = 0$, $u_{n^{(k)}} = \delta$ $(u_{n_1^{(k_1)}} = -\delta$, $u_{n_1^{(k_1)}} = 0$, $u_{n_1^{(k_1)}} = \delta)$ respectively. Obviously, B (B_1) lies between A (A_1) and C (C_1). A point moving on l from the α- to the ω-end-point of l conceivably meets A, C, A_1, C_1 in five possible orders, namely,

$$(30) \qquad ACA_1C_1, \quad AA_1CC_1, \quad AA_1C_1C, \quad A_1ACC_1, \quad A_1AC_1C.$$

If we take these cases in the order given, the directed H-line segments

$$(31) \qquad \overrightarrow{AC_1}, \quad \overrightarrow{AC_1}, \quad \overrightarrow{AC}, \quad \overrightarrow{A_1C_1}, \quad \overrightarrow{A_1C}$$

intersect $\Delta_{n^{(k)}}$ and $\Delta_{n_1^{(k_1)}}$. Consider the region

$$(32) \qquad -3\delta < u_{n^{(k)}} \leqq -\delta, \qquad 0 \leqq v_{n^{(k)}} \leqq |n'|$$

of Φ. The α-end-points of the directed H-line segments in (31) all lie in (32). This statement is trivial for the first three cases. For the remaining two cases we note that $|A_1B| < |A_1B_1|$. According to Lemma 6, $|A_1B_1| < |n_1'| + \delta$ and therefore, since

$$(33) \qquad |n_1'| < 2\delta,$$

inasmuch as $\sinh(|n_1'|/2) < \tanh\delta < \sinh\delta$, we have $|A_1B| < 3\delta$, which proves the statement for the remaining two cases.

A set $S_{n^{(k)}}$ of directed H-line segments having property I therefore exists such that the α-end-point of each member lies on (22).

Let s be an arbitrary member of such a set $S_{n^{(k)}}$. We shall show that $|s|$ can be taken less than λ_1, where λ_1 is defined in (23). Denote the α-end-point of s by α.

For the first case in (30), construct the broken H-line segment $\alpha A'C_1'C_1$, where A' is a point on the side e_k of D_0 along which $\bar{D}_0$, $\bar{D}_k$ border and C_1' is a point on the side e_{k_1} of D_0 along which $\bar{D}_0$, $\bar{D}_{k_1}$ border. Since $\bar{D}_k$ is contained in the region [see (20)]

$$T_{n^{(k)}}: \qquad 0 < u_{n^{(k)}} \leqq \tau_{n'}, \qquad 0 \leqq v_{n^{(k)}} \leqq |n'|,$$

and $\bar{D}_{k_1}$ is contained in the region [see (21)]

$$\overline{T}_{n_1^{(k_1)}}: \qquad -\tau_{n_1'} \leqq u_{n_1^{(k_1)}} < 0, \qquad 0 \leqq v_{n_1^{(k_1)}} \leqq |n_1'|,$$

we have, on using Lemma 6,

$$|\alpha A'| < |n'| + 3\delta + \tau_{n'}, \qquad |C_1'C_1| < |n_1'| + \delta + \tau_{n_1'}.$$

Now $|A'C_1'| \leqq \delta$, so that

$$(34) \qquad |\alpha C_1| < 5\delta + |n'| + |n_1'| + \tau_{n'} + \tau_{n_1'},$$

inasmuch as $|\alpha C_1| \leqq |\alpha A'| + |A'C_1'| + |C_1'C_1|$. Using (33) and the inequality

$$|n'| < 2\delta, \tag{35}$$

which is proved in the same way as (33), and employing (7), we find

$$|\alpha C_1| < 11\delta + \log \frac{(4 \sinh \delta)^2}{|n'|\,|n_1'|}. \tag{36}$$

Before disposing of the remaining four cases, we note that

$$|AC| = 2\delta, \qquad |A_1C_1| < |n_1'| + 2\delta < 4\delta.$$

For the second case, we have

$$|\alpha C_1| \leqq |\alpha C| + |A_1C_1| < 4\delta + 4\delta = 8\delta, \tag{37}$$

inasmuch as $|\alpha C| = |\alpha B| + |BC| = 4\delta$.

In the third case, we have, since $|\alpha A| = 2\delta$,

$$|\alpha C| = |\alpha A| + |AC| = 2\delta + 2\delta = 4\delta. \tag{38}$$

In the fourth case, we have, since $|\alpha A_1| \leqq 2\delta$,

$$|\alpha C_1| = |\alpha A_1| + |A_1C_1| < 2\delta + 4\delta = 6\delta. \tag{39}$$

For the fifth case, we proceed as in the third case, and obtain the same inequality (38).

On comparing (36), (37), (38), (39) with (23), we see that $|s|$ can always be taken less than λ_1.

Property III. Consider the region

$$R_1: \qquad -\infty < u_{n_1}^{(k_1)} < +\infty, \qquad 0 \leqq v_{n_1}^{(k_1)} \leqq |n_1'|$$

of Φ and denote the two remaining regions of Φ by R_2, R_3. From (28), we see that $n^{(k)}$ lies in R_1. Those members of $F_{n^{(k)}}$ which intersect (29) (and therefore intersect $\Delta_{n_1}^{(k_1)}$) intersect $n^{(k)}$ in the points of an arc μ_1, the two members of $F_{n^{(k)}}$ drawn through the end-points of μ_1 passing through the end-points of (29) into the regions R_2, R_3. Denote these two members of $F_{n^{(k)}}$ by r_2, r_3, choosing the notation so that r_2 enters the region R_2. A point displaced along r_2 (r_3) from $n^{(k)}$ to the ω-end-point of r_2 (r_3) enters, and never leaves R_2 (R_3).

One readily sees from the analysis accompanying property II that when $v_{n^{(k)}} =$ const. intersects (29), the H-length of the segment of it comprehended between $n^{(k)}$ and (29) is less than λ_1. Hence $u_{n^{(k)}} = \lambda_1$ intersects r_2, r_3 at points interior to R_2, R_3 respectively. Therefore the H-length of the segment of $u_{n^{(k)}} = \lambda_1$ comprehended between r_2 and r_3 exceeds $|n_1'|$, since $|n_1'|$ is the greatest lower bound of the H-distances between points of R_2 and R_3.

On placing $n = \mu_1$, $k = \lambda_1$, $\kappa > |n_1'|$ in (2), we find

$$|\mu_1| \cosh \lambda_1 > |n_1'|,$$

and therefore that

$$|\mu_1| > \frac{2|n_1'|}{e^{\lambda_1} + e^{-\lambda_1}} > |n_1'|e^{-\lambda_1}.$$

A set $S_{n^{(k)}}$ of directed H-line segments having properties I and II of the lemma, and which also possesses property III, therefore aways exists.

Lemma 21 forms the basis for the following lemma.

LEMMA 22. *Let q denote a positive integer and let $\Delta_{n'}, \Delta_{n_1'}, \Delta_{n_2'}, \cdots, \Delta_{n_q'}$ be $1 + q$ δ-sets based on the directed H-line segments $n', n_1', n_2', \cdots, n_q'$ respectively, where*

$$\text{(40)} \quad \sinh(|n'|/2) < \tanh\delta, \qquad \sinh(|n_j'|/2) < \tanh\delta \qquad (j = 1, 2, \cdots, q).$$

Let $F_{n'}$ be the orthogonal family of directed H-lines based on n' and associated with $\Delta_{n'}$.

Among the copies of $F_{n'}$ there is one, say $F_{n^{[q]}}$, which contains a set $S_{n^{[q]}}$ of directed H-line segments having the following properties:

I. *The members of $S_{n^{[q]}}$ intersect the copy $\Delta_{n^{[q]}}$ of $\Delta_{n'}$ and intersect copies of each of the δ-sets $\Delta_{n_1'}, \Delta_{n_2'}, \cdots, \Delta_{n_q'}$.*

II. *The α-end-points of the members of $S_{n^{[q]}}$ lie on the H-line segment*

$$\text{(41)} \qquad u_{n^{[q]}} = -3\delta, \qquad 0 \leqq v_{n^{[q]}} \leqq |n'|,$$

and the H-length of each member is less than λ_q, where

$$\text{(42)} \qquad \lambda_q = 11\delta q + \log \frac{(4 \sinh \delta)^{2q}}{|n'|\,|n_1'|^2\,|n_2'|^2 \cdots |n_{q-1}'|^2\,|n_q'|}.$$

III. *The members of $S_{n^{[q]}}$ intersect $n^{[q]}$ in the points of an arc μ_q which contains an arc ν_q' whose H-length is given by*

$$\text{(43)} \qquad |\nu_q'| = |n_q'|e^{-\lambda_q}.$$

For $q = 1$, this lemma is equivalent to Lemma 21. The proof for an arbitrary q is now obtained by induction.

Suppose the lemma is true for $q = r - 1$. On this assumption, there is a copy $F_{n^{[r-1]}}$ of $F_{n'}$ which contains a set $S_{n^{[r-1]}}$ of directed H-line segments having the properties I, II, III of the lemma, q being replaced by $r - 1$.

Let $F_{\nu_{r-1}'}$ denote the orthogonal family of directed H-lines composed of those members of $F_{n^{[r-1]}}$ which intersect ν_{r-1}' and let $S_{\nu_{r-1}'}$ denote the subset of $S_{n^{[r-1]}}$ composed of those members of $S_{n^{[r-1]}}$ which intersect ν_{r-1}'. Let $\Delta_{\nu_{r-1}'}$ denote the δ-set based on ν_{r-1}' and associated with $F_{\nu_{r-1}'}$. We have

$$\text{(44)} \qquad \Delta_{n^{[r-1]}} \supset \Delta_{\nu_{r-1}'},$$

and note that $S_{\nu_{r-1}'}$ possesses the properties I, II, III ($q = r - 1$) possessed by $S_{n^{[r-1]}}$.

Apply Lemma 21 to $\Delta_{\nu_{r-1}'}, \Delta_{n_r'}$. Using the same notation as was employed in §7 and Lemma 21, we see that among the $F_{\nu_{r-1}'}^{(i)}$ $(i = 1, 2, \cdots, 2p)$ there is

one, say $F_{\nu_{r-1}^{(k)}}$, which contains a set $S_{\nu_{r-1}^{(k)}}$ of directed H-line segments having the following properties:

I′. Each member of $S_{\nu_{r-1}^{(k)}}$ intersects $\Delta_{\nu_{r-1}^{(k)}}$ and there exists a $\Delta_{n_r^{(i)}}$, say $\Delta_{n_r^{(k_1)}}$, which is intersected by all the members of $S_{\nu_{r-1}^{(k)}}$.

II′. The α-end-points of the members of $S_{\nu_{r-1}^{(k)}}$ lie on the H-line segment

$$u_{\nu_{r-1}^{(k)}} = -3\delta, \qquad 0 \leqq v_{\nu_{r-1}^{(k)}} \leqq |\nu'_{r-1}|, \tag{45}$$

and the H-length of each member is less than $\bar{\lambda}_r$, where

$$\bar{\lambda}_r = 11\delta + \log \frac{(4 \sinh \delta)^2}{|\nu'_{r-1}|\,|n'_r|}. \tag{46}$$

III′. The members of $S_{\nu_{r-1}^{(k)}}$ intersect $\nu_{r-1}^{(k)}$ in the points of an arc $\bar{\mu}_r$ which contains an arc $\bar{\nu}'_r$ whose H-length is given by

$$|\bar{\nu}'_r| = |n'_r| e^{-\bar{\lambda}_r}. \tag{47}$$

Placing $q = r - 1$ in (43) and substituting the resulting expression for $|\nu'_{r-1}|$ in (46), we find

$$\bar{\lambda}_r = 11\delta + \lambda_{r-1} + \log \frac{(4 \sinh \delta)^2}{|n'_{r-1}|\,|n'_r|},$$

and therefore, from (42) for $q = r - 1$, we obtain

$$\bar{\lambda}_r = 11\delta_r + \log \frac{(4 \sinh \delta)^{2r}}{|n'|\,|n'_1|^2\,|n'_2|^2 \cdots |n'_{r-1}|^2\,|n'_r|},$$

so that

$$\bar{\lambda}_r = \lambda_r. \tag{48}$$

On comparing (47) and (48) with (43), we find

$$|\bar{\nu}'_r| = |\nu'_r|. \tag{49}$$

Using (40), we easily show that

$$\lambda_{r-1} < \lambda_r. \tag{50}$$

Let $\Sigma_{\nu_{r-1}^{(k)}}$ denote the transform of $S_{\nu'_{r-1}}$ by the transformation of G which carries $F_{\nu'_{r-1}}$ into $F_{\nu_{r-1}^{(k)}}$. $\Sigma_{\nu_{r-1}^{(k)}}$ has the following properties:

I″. The members of $\Sigma_{\nu_{r-1}^{(k)}}$ intersect the copy $\Delta_{\nu_{r-1}^{(k)}}$ of $\Delta_{\nu'_{r-1}}$ and intersect copies of the δ-sets $\Delta_{n'_1}, \Delta_{n'_2}, \cdots, \Delta_{n'_{r-1}}$.

II″. The α-end-points of the members of $\Sigma_{\nu_{r-1}^{(k)}}$ fill up the H-line segment (45) and the H-length of each member is less than λ_{r-1}.

III″. The intersection points of the members of $\Sigma_{\nu_{r-1}^{(k)}}$ with $\nu_{r-1}^{(k)}$ fill up $\nu_{r-1}^{(k)}$.

According to I′, I″, the members of both $S_{\nu_{r-1}^{(k)}}$, $\Sigma_{\nu_{r-1}^{(k)}}$ intersect a copy of $\Delta_{\nu'_{r-1}}$. From (44), they therefore intersect a copy of $\Delta_{n'}$.

In order to construct a set $S_{n^{[r]}}$ of directed H-line segments possessing the properties I, II, III for $q = r$, we take $n^{[r]}$ as that copy of n' which contains

$\nu_{r-1}^{(k)}$ as a segment. $F_{n[r]}$ is the copy of $F_{n'}$ based on $n^{[r]}$. On comparing I′, II′, III′ with I″, II″, III″ in the light of (48), (49), (50), we see that a set $S_{n[r]}$ exists; the arc ν_r' in III ($q = r$) being taken as the arc $\bar{\nu}_r'$ in III′, and the members of $S_{n[r]}$ constructed as directed H-line segments drawn perpendicular to $\bar{\nu}_r'$ with their α-end-points on (45) and their H-lengths less than λ_r.

The following corollary is important in the subsequent development of the paper.

COROLLARY. *If the H-lengths* $|n'|, |n_1'|, |n_2'|, \cdots, |n_q'|$ *all exceed a positive constant* γ, *the H-length of each member of* $S_{n[q]}$ *is less than* λ_q, *where*

$$\lambda_q < q\left[11\delta + 2\log\frac{4\sinh\delta}{\gamma}\right].$$

The proof of this corollary is immediate in view of the value given for λ_q in property II of the above lemma.

9. **The surface φ and the phase space M.** When congruent points on Φ are taken as identical, a topologically closed, orientable surface ϕ of genus p and of constant curvature -1 is obtained.[15]

Let g_Φ denote an arbitrary directed H-line segment on Φ. From Lemmas 15 and 17 it follows that g_Φ is divided into a finite number of directed segments by the meshes of N, each segment being directed in the same sense as g_Φ. If a mesh containing a segment of g_Φ is transformed into $\bar{D}_0$ by a properly chosen transformation of G, the transform of the segment of g_Φ which it contains is a directed H-line segment lying in $\bar{D}_0$. The totality of directed H-line segments lying in $\bar{D}_0$ obtained in this manner from g_Φ constitutes a *directed geodesic segment* g_φ *on* φ. g_Φ *represents* g_φ on Φ. This representation is not unique, any copy of g_Φ also representing g_φ on Φ. The directed H-line segments in $\bar{D}_0$ comprising g_φ are *pieces of* g_φ. The sum of the H-lengths of the pieces of g_φ is the *H-length of* g_φ and equals the H-length of g_Φ.

Let P be an arbitrary point on Φ and to P attach a direction θ measured in the usual way in the x, y-plane. Following Morse, the pair (P, θ) is an *element.* Two elements whose angles differ by an integral multiple of 2π will be taken as identical.[16] An element (P, θ) is *on* a directed H-line segment g_Φ if P is a point of g_Φ and θ is the direction of g_Φ at P. Two elements (P', θ'), (P'', θ'') are *copies* of one another if P' can be transformed into P'' by a transformation of G which takes θ' into θ''. The *distance between two elements* (P', θ'), (P'', θ'') will be defined to be

$$|P'P''| + \min|\theta' - \theta'' + 2n\pi|;$$

[15] See, for example, Morse II or Nielsen, loc. cit.

[16] Morse I, p. 52. Our convention differs from that of Morse, since he regards two directions as identical when they differ by an integral multiple of π, whereas we take two directions identical if they differ by an integral multiple of 2π.

where min $| \theta' - \theta'' + 2n\pi |$ represents the minimum of $| \theta' - \theta'' + 2n\pi |$ for all integers n, positive, negative or zero.[17] A pair of elements (P_0', θ'), (P_0'', θ'') in which P_0', P_0'' are points of $\bar{D}_0$ possess a φ-distance equal to the greatest lower bound of the distances from the copies of one to the copies of the other.[17]

The set of elements (P_0, θ) for which P_0 lies in $\bar{D}_0$ constitutes the *phase space* M. The *distance between two elements in* M is taken to be their φ-distance as defined above. The set of elements in M afforded by the points and directions on a directed geodesic segment g_φ is a *phase curve* g_M, and we shall say g_φ *generates* g_M. If g_Φ represents g_φ on Φ, we shall also say that g_Φ generates g_M.

10. **The subsets $S_M^{(1)}$, S_M, S_M^* of M.** Let g_Φ be a directed H-line segment on Φ which terminates on N, intersects $\bar{D}_0$, and meets no vertex of N. In addition g_Φ is supposed to have an admissible sequence

$$a_1 a_2 \cdots a_m a_{m+1} \cdots a_{2m}, \tag{51}$$

in which a_m, a_{m+1} arise from the intersections of g_Φ with the boundary of D_0. Consider the set of directed H-line segments such as g_Φ which have their α (ω)-end-points on the side a_1 (a_{2m}) of N. Denote this set by $S^{(1)}$, and consider the subset of $S^{(1)}$ formed by those directed H-line segments which are segments of members of $S^{(1)}$ and which lie on $\bar{D}_0$. Denote this subset by $S_{\bar{D}_0}^{(1)}$ and let $S_M^{(1)}$ denote the set of elements in M which are on the members of $S_{\bar{D}_0}^{(1)}$. $S^{(1)}$, $S_{\bar{D}_0}^{(1)}$, $S_M^{(1)}$ are *generated* by the admissible sequence (51).

For fixed m, the number of admissible sequences (51) cannot exceed $(4p)^{2m}$, so that the number of sets $S_M^{(1)}$ does not exceed $(4p)^{2m}$. An arbitrary element (P_0, θ) of M is, however, contained in at least one $S_M^{(1)}$. For take the point P_0 of $\bar{D}_0$ and construct the directed H-line $\overrightarrow{AB}$ passing through P_0 with the direction θ. If $\overrightarrow{AB}$ does not meet a vertex of N, a segment $\overrightarrow{A_1B_1}$ of $\overrightarrow{AB}$ can be taken which has an admissible sequence such as (51), and the set $S_M^{(1)}$ generated by this admissible sequence contains (P_0, θ). When $\overrightarrow{AB}$ meets one or more vertices of N two cases are possible: (a) $\overrightarrow{AB}$ coincides with an H-line in N; (b) $\overrightarrow{AB}$ does not coincide with an H-line in N.

Let us consider (a). Here $\overrightarrow{AB}$ is divided into segments each of H-length ρ by the vertices of N and from any vertex V of N on $\overrightarrow{AB}$ radiate $2p - 1$ H-rays of N into each of the regions Φ_1, Φ_2 into which Φ is divided by AB. The $2p - 1$ H-rays radiating from V into Φ_1 are now directed positively away from V. A directed H-ray r_1 *precedes* a directed H-ray r_2 if the angle at V taken positively and less than π between r_1 and the *negative* sense on $\overrightarrow{AB}$ is less than the corresponding angle for r_2. Giving those directed H-rays emanating from a vertex $V^{(1)}$ on $\overrightarrow{AB}$ precedence over those emanating from a vertex $V^{(2)}$ of $\overrightarrow{AB}$ if $V^{(1)}$ precedes $V^{(2)}$ on $\overrightarrow{AB}$, we may arrange the directed H-rays radiating from the vertices of N on $\overrightarrow{AB}$ into Φ_1 in an unending sequence

$$\cdots r_{-m} r_{-m+1} \cdots r_{-2} r_{-1} r_1 r_2 \cdots r_{m-1} r_m \cdots. \tag{52}$$

[17] Compare Morse I, p. 53.

Let us take Φ_1 to be the region of Φ containing D_0 and choose the notation in the above sequence so that r_{-1}, r_1 correspond to the sides of D_0. Consider the set $S^{(1)}$ comprising those directed H-line segments with their α (ω)-end-points on that side of N which lies on r_{-m} (r_m) and meets $\overrightarrow{AB}$. The element (P_0, θ) is on the segment of $\overrightarrow{AB}$ which belongs to $S^{(1)}$. Consequently (P_0, θ) is contained in a set $S_M^{(1)}$.

The treatment of (b) requires a slight modification of the procedure in (a). Here $\overrightarrow{AB}$ meets N at points of two types, those which are vertices of N and those which are not vertices of N. As before, Φ is divided by AB into two regions Φ_1, Φ_2, each of which may now contain points of D_0. In certain instances one of Φ_1, Φ_2 may contain no points of D_0, as is the case when AB meets D_0 in a vertex only. In any event, Φ_1 shall denote one of the regions which contains points of D_0. When $\overrightarrow{AB}$ meets a vertex V of N, exactly $2p$ H-rays in N radiate from V into Φ_1. These $2p$ H-rays are then directed and ordered by a repetition of the process employed in (a) for the $2p - 1$ H-rays there considered. When $\overrightarrow{AB}$ meets a point P of N which is not a vertex of N, only one H-ray in N radiates from P into Φ_1. This H-ray is directed positively from P. A convention similar to that employed above in the construction of the unending sequence (52) then leads us to a similar sequence, and from this point onward the procedure is analogous to that in (a).

Take $m = 2pr + 1$ in (51), where

$$r > \rho/\chi \tag{53}$$

and is a positive integer. From Lemma 15, we see that

$$|[a_1 a_m]| \geqq r\chi > \rho, \qquad |[a_{m+1} a_{2m}]| \geqq r\chi > \rho,$$

and therefore, since $|g_\Phi| > |[a_1 a_m]| + |[a_{m+1} a_{2m}]|$, we have

$$|g_\Phi| > 2\rho. \tag{54}$$

This last inequality shows that a_1, a_{2m} cannot have a common end-point. For, if they do, g_Φ lies in the H-circle of H-radius ρ described about the common end-point as an H-center, so that $|g_\Phi| \leqq 2\rho$, which contradicts (54). In addition, a_1, a_{2m} cannot be segments of the same H-line of N, as this would require g_Φ to meet a vertex of N.

Four distinct members of $S^{(1)}$ therefore exist which connect end-points of a_1 to end-points of a_{2m}. Two of these, say l, l', do not intersect. At least one, say l, intersects the boundary of D_0. For suppose this is not the case. Since g_Φ intersects the boundary of D_0, the interior of the region of Φ enclosed by l, l', a_1, a_{2m} contains D_0 together with its boundary, and hence contains the H-circle (12) of H-radius $\rho/2$. This last is impossible, according to Lemma 4; thus l may be taken as intersecting the boundary of D_0.

Let A, B (A_0, B_0) denote the α- and ω-end-points respectively of l (g_Φ). Let A', B' denote the points (possibly coincident) where l intersects[18] the boundary

[18] In case l intersects the boundary of D_0 in a side of D_0, A', B' denote the end-points of this side.

of D_0, choosing the notation so that $|AA'| \leqq |AB'|$. A review of the proof of Lemma 17 shows that a member $\overrightarrow{A_1B_1}$ of $S^{(1)}$ not meeting a vertex of N can be selected so that it intersects the boundary of D_0 in points A_1', B_1' which lie on the same sides of D_0 as do the points A', B' respectively, and in addition is such that A_1 (B_1) lies on a_1 (a_{2m}) between $A(B)$ and A_0 (B_0). Using the method employed in the proof of Lemma 16 to deform $\overrightarrow{A_0B_0}$ into $\overrightarrow{A_1B_1}$, we easily see that the number of symbols in the admissible sequences of $\overrightarrow{A_1A_1'}$ and $\overrightarrow{B_1'B_1}$ is in each case equal to $m = 2pr + 1$. On interpreting $\overrightarrow{A_1B_1}$ in Lemma 17 as $\overrightarrow{AA'}$ $(\overrightarrow{B'B})$ and $\overrightarrow{AB}$ in Lemma 15 as $\overrightarrow{A_1A_1'}$ $(\overrightarrow{B_1'B_1})$, we have

$$(55) \qquad |AA'| \geqq r\chi > \rho, \qquad\qquad |B'B| \geqq r\chi > \rho,$$

so that $|AB| > 2\rho$, and therefore

$$|AB| = 2\lambda_1 + 2\rho \qquad\qquad (\lambda_1 > 0).$$

We introduce two Gaussian geodesic coördinate systems $[u_n, v_n]$, $[u_{n'}, v_{n'}]$. In the first, A, B are the points $[-(\lambda_1 + \rho), 0]$ and $[\lambda_1 + \rho, 0]$ respectively, and the region $v_n \geqq 0$ of Φ contains the region of Φ bounded by l, l', a_1, a_{2m}. In the second, A', B' are the points $[-k, 0]$, $[k, 0]$ respectively (k is a constant, positive or zero) and the region $v_{n'} \geqq 0$ of Φ contains the region of Φ bounded by l, l', a_1, a_{2m}.

The points $R'[-r\chi, 0]$, $R''[r\chi, 0]$ in the $[u_{n'}, v_{n'}]$ coördinate system lie on $\overrightarrow{AB}$. This follows from (55). Let C_1', C', C'', C_1'' denote the H-semicircles drawn in the region $v_{n'} \geqq 0$ of Φ about the points A, R', R'', B as H-centers respectively, each H-semicircle having an H-radius ρ. Let S denote the set of directed H-lines which intersect both C', C'' and are directed from C' to C''. The members of S intersect $u_{n'} = 0$ in the points of an H-line segment which we denote by n'. n' is therefore the H-line segment

$$(56) \qquad u_{n'} = 0, \qquad\qquad 0 \leqq v_{n'} \leqq |n'|,$$

and $v_n = |n'|$ is tangent to both C', C''. Since a_1 (a_{2m}) is an H-radius of C_1' (C_1''), one perceives, on identifying the $[u_n, v_n]$ coördinate system introduced above with the $[u_n, v_n]$ coördinate system in Lemma 3, that

$$(57) \qquad S \supset S^{(1)}.$$

If $S_{\bar{D}_0}$ denotes the subset of S which is formed by directed H-line segments which are segments of members of S and which lie in $\bar{D}_0$, we have, from (57), the fact that $S_{\bar{D}_0} \supset S_{\bar{D}_0}^{(1)}$. Hence if S_M denotes the set of elements in M which are on the members of $S_{\bar{D}_0}$, we see that $S_M \supset S_M^{(1)}$.

For a fixed m, the number of sets S_M does not exceed $(4p)^{2m}$, each element of M being contained in at least one set S_M.

Next we note that $S_{\bar{D}_0}$ is a subset of a δ-set $\Delta_{n'}$ based on n', the α (ω)-end-points of the members of $\Delta_{n'}$ composing the points of the arc $0 \leqq v_{n'} \leqq |n'|$ of $u_{n'} = -\delta$ $(u_{n'} = \delta)$. We observe that if ξ denotes the H-distance from the center of Ψ to $u_{n'} = 0$, we have $\xi \leqq \delta/2 < \delta$. Those elements which are on

members of $\Delta_{n'}$ and belong to M constitute a set S_M^* of elements in M which contains S_M.

For a given m, the number of sets S_M^* does not exceed $(4p)^{2m}$, each element of M belonging to at least one S_M^*.

11. **The function** $n(\epsilon)$. Consider a δ-set $\Delta_{n'}$ of directed H-line segments based on a directed H-line segment n' which is a segment of a directed H-line whose H-distance ξ from the center of Ψ is less than δ. Choose a member of $\Delta_{n'}$ at random. Let ϵ denote an arbitrary, preassigned positive number. We propose to calculate a function $n(\epsilon)$ of ϵ such that if $|n'| \leqq n(\epsilon)$, an arbitrary element on an arbitrary member of $\Delta_{n'}$ lies within a distance ϵ of some element on the member of $\Delta_{n'}$ chosen at random above.

We begin by introducing the Gaussian geodesic coördinate system $[u_{n'}, v_{n'}]$ on Φ associated with the directed H-line segment n'. The α (ω)-end-points of the members of $\Delta_{n'}$ lie on $u_{n'} = -\delta$ ($u_{n'} = \delta$). Let (P, θ) be an arbitrary element on an arbitrary member of $\Delta_{n'}$, and let $[a, b]$ denote the coördinates of P in the $[u_{n'}, v_{n'}]$ coördinate system. The member of $\Delta_{n'}$ chosen at random above intersects $u_{n'} = a$ at a point $P'[a, b']$ and has a direction θ' at this point. We proceed to derive an upper bound for the distance between the elements (P, θ), (P', θ').

The H-distance $|PP'|$ cannot exceed the H-length of the segment of $u_{n'} = a$ lying between P and P', and this H-length cannot exceed the H-length of the segment of $u_{n'} = a$ comprehended between $v_{n'} = 0$, $v_{n'} = |n'|$, so that from (2) we find

$$|PP'| \leqq |n'| \cosh a. \tag{58}$$

Now $-\delta \leqq a \leqq \delta$. Therefore

$$\cosh a \leqq \cosh \delta, \tag{59}$$

and hence

$$|PP'| < |n'| \cosh \delta, \tag{60}$$

since the equality signs in (58) and (59) cannot hold simultaneously.

We now take up $|\theta - \theta'|$. Let Q denote an arbitrary point of the region

$$-\delta \leqq u_{n'} \leqq \delta, \qquad 0 \leqq v_{n'} \leqq |n'| \tag{61}$$

of Φ. Denote by ζ (ζ') the magnitude of the angle filled up at Q by the directions of those H-rays drawn from Q to intersect the segment $0 \leqq v_{n'} \leqq |n'|$ of $u_{n'} = \delta$ ($u_{n'} = -\delta$). If $[u_{n'}, v_{n'}]$ are the coördinates of Q, the angles ζ, ζ' are functions of $u_{n'}$, $v_{n'}$, and we write[19]

$$\zeta = \zeta(u_{n'}, v_{n'}), \qquad \zeta' = \zeta'(u_{n'}, v_{n'}).$$

[19] Here $\zeta(\delta, v_{n'})$, $\zeta'(-\delta, v_{n'})$ are not properly defined. We place

$$\zeta(\delta, v_{n'}) \equiv \zeta'(-\delta, v_{n'}) \equiv 2\pi.$$

If we define $Z = \min\{\zeta, \zeta'\}$, we have $Z = Z(u_{n'}, v_{n'})$ yielding an upper bound for the magnitude of the angle filled up at Q by the directions of those members of $\Delta_{n'}$ which pass through Q. If we hold $v_{n'}$ fast, ζ (ζ') is seen to be a monotone increasing (decreasing) function of $u_{n'}$ for $-\delta \leqq u_{n'} \leqq \delta$. Moreover, from reasons of symmetry, it is apparent that $\zeta(-u_{n'}, v_{n'}) = \zeta'(u_{n'}, v_{n'})$, so that $\zeta(0, v_{n'}) = \zeta'(0, v_{n'})$. Hence

$$Z = \zeta \text{ when } -\delta \leqq u_{n'} \leqq 0, \qquad Z = \zeta' \text{ when } 0 \leqq u_{n'} \leqq \delta,$$

and therefore, for a given $v_{n'}$,

$$Z(u_{n'}, v_{n'}) \leqq Z(0, v_{n'}), \qquad -\delta \leqq u_{n'} \leqq \delta.$$

Finally, an elementary calculation which is rather tedious and need not be given here shows that[20]

$$Z(0, v_{n'}) \leqq Z(0, |\,n'\,|/2), \qquad 0 \leqq v_{n'} \leqq |\,n'\,|,$$

so that

$$Z(u_{n'}, v_{n'}) \leqq Z(0, |\,n'\,|/2), \qquad -\delta \leqq u_{n'} \leqq \delta, \qquad 0 \leqq v_{n'} \leqq |\,n'\,|. \tag{62}$$

In order to obtain an upper bound for $Z(0, |\,n'\,|/2)$, let us introduce a system of geodesic polar coördinates $[r, \psi]$ on Φ, taking the pole as the point $[0, |\,n'\,|/2]$ in the $[u_{n'}, v_{n'}]$ coördinate system and measuring ψ positively in the counter-clockwise sense from the coördinate line $\psi = 0$, which is taken as the H-ray drawn from the pole to the point with coördinates $[\delta, 0]$ in the $[u_{n'}, v_{n'}]$ coördinate system. According to (2), the H-length of the segment of $u_{n'} = \delta$ which is comprehended between $v_{n'} = 0$, $v_{n'} = |\,n'\,|$ is $|\,n'\,| \cosh \delta$. On using (P) to recalculate this H-length, one finds

$$|\,n'\,| \cosh \delta = \int \{dr^2 + \sinh^2 r d\psi^2\}^{\frac{1}{2}}.$$

Comparing the circle $r = \delta$ with the circle of which $u_{n'} = \delta$ is a segment, we have $r \geqq \delta$ in the above integration, so that

$$|\,n'\,| \cosh \delta > \int \{dr^2 + \sinh^2 \delta d\psi^2\}^{\frac{1}{2}} > \int_0^{Z(0, |n'|/2)} \sinh \delta d\psi,$$

from which we establish that

$$Z(0, |\,n'\,|/2) < |\,n'\,| \coth \delta. \tag{63}$$

On combining (62), (63), we see that the magnitude of the angle filled up at an arbitrary point Q of the region (61) by the directions of the members of $\Delta_{n'}$ passing through Q is less than $|\,n'\,| \coth \delta$.

We are now in a position to obtain an upper bound for $|\,\theta - \theta'\,|$. Take the member of $\Delta_{n'}$ passing through P (P') which is a segment of the coördinate line

[20] Professor Morley has kindly pointed out to me that this result is obtained very neatly by using the methods of inversive geometry.

$v_{n'} = b$ ($v_{n'} = b'$), directed in the sense of increasing $u_{n'}$, and denote its direction at P (P') by α (α'). We have

$$| \theta - \theta' | \leqq | \theta - \alpha | + | \alpha - \alpha' | + | \alpha' - \theta' |.$$

Now neither of $| \theta - \alpha |$, $| \alpha' - \theta' |$ can exceed $| n' | \coth \delta$, and from (3) we have $| \alpha - \alpha' | < | n' | \sinh \delta$, inasmuch as $\xi < \delta$. Hence

$$| \theta - \theta' | < | n' | \{2 \coth \delta + \sinh \delta\}. \tag{64}$$

From (60), (64) we have

$$| PP' | + | \theta - \theta' | < | n' | \{\cosh \delta + 2 \coth \delta + \sinh \delta\}, \tag{65}$$

so that, on referring to the definition for the distance between (P, θ), (P', θ') given in §9, we see that the function $n(\epsilon)$ defined by

$$n(\epsilon) = \epsilon\{\cosh \delta + 2 \coth \delta + \sinh \delta\}^{-1} \tag{66}$$

possesses the desired property.

12. **The upper bound for the ergodic function.** Let ϵ denote a positive number satisfying the inequalities[21]

$$\epsilon < (\cosh \delta + 2 \coth \delta + \sinh \delta)(1 - e^{-2\rho}), \tag{67}$$

$$\epsilon < 2(\cosh \delta + 2 \coth \delta + \sinh \delta) \log\{\tanh \delta + (1 + \tanh^2 \delta)^{\frac{1}{2}}\}, \tag{68}$$

where ρ, δ are defined in (10), (11) respectively, and let r denote the positive integer bounded by the inequalities

$$\frac{1}{\chi} \log \frac{2 \sinh \rho}{n(\epsilon)} \leqq r < 1 + \frac{1}{\chi} \log \frac{2 \sinh \rho}{n(\epsilon)}. \tag{69}$$

Inequality (53) follows from (67), (66) and the lower bound in (69). From (68) we find

$$\sinh \frac{n(\epsilon)}{2} < \tanh \delta. \tag{70}$$

Going back to the construction of n' in (56) and using Lemma 5, we see that

$$\rho e^{-r\chi} < | n' | < 2e^{-r\chi} \sinh \rho, \tag{71}$$

when we place $\lambda + \rho = r\chi$ in (6). On using (69), we find that (71) is replaced by

$$\frac{\rho}{2e^{\chi} \sinh \rho} n(\epsilon) < | n' | < n(\epsilon). \tag{72}$$

Now consider a subset S_M^* of M and suppose the δ-set from which it is obtained to be based on n'. The upper bound in (72) coupled with the results in

[21] One of these inequalities probably insures the holding of the other. We need not however, determine which inequality possesses this property.

§11 shows that the phase curve generated in M by an arbitrary member of $\Delta_{n'}$ comes within a distance ϵ of every point of S_M^*, the distance between elements of M being defined as in §9.

r being chosen subject to (69), the number $2m$ ($m = 2pr + 1$) of symbols in (51) is consequently bounded by the inequalities

$$2 + \frac{4p}{\chi} \log \frac{2 \sinh \rho}{n(\epsilon)} \leqq 2m < 4p + 2 + \frac{4p}{\chi} \log \frac{2 \sinh \rho}{n(\epsilon)}. \tag{73}$$

Let $1 + q$ denote the number of sets S_M^* occurring when m is chosen as above. Since $1 + q < (4p)^{2m}$, we have

$$1 + q < (4p)^{4p+2}(4p)^{\frac{4p}{\chi} \log \frac{2 \sinh \rho}{n(\epsilon)}},$$

which reduces to

$$1 + q < (4p)^{4p+2} \left(\frac{2 \sinh \rho}{n(\epsilon)}\right)^{\frac{4p \log 4p}{\chi}}. \tag{74}$$

From the upper bound in (72) and the inequality (70), it is seen that Lemma 22 may be applied to the $1 + q$ δ-sets introduced above. If we apply Lemma 22, it follows that a directed H-line segment g_Φ of H-length less than λ_q exists which intersects copies of these $1 + q$ δ-sets. From the lower bound in (72) we see that we can set $\gamma = \frac{1}{2}\rho e^{-\chi} n(\epsilon) \operatorname{csch} \rho$ in the upper bound for λ_q as given in the corollary to Lemma 22. On doing this and making use of (74), we find

$$\lambda_q < \left[(4p)^{4p+2}\left(\frac{2 \sinh \rho}{n(\epsilon)}\right)^{\frac{4p \log 4p}{\chi}} - 1\right]\left[2 \log \frac{8e^\chi \sinh \delta \sinh \rho}{\rho n(\epsilon)} + 11\delta\right]. \tag{75}$$

If the -1 in (75) is repressed[22] and if $n(\epsilon)$ is replaced by its value given in (66), the inequality (75) may be replaced by one of the form

$$\lambda_q < \epsilon^{-\omega}\left[A \log \frac{B}{\epsilon} + C\right], \tag{76}$$

where A, B, C and ω are given in terms of p as follows:

$$A = 2(4p)^{4p+2}\{2 \sinh \rho(\cosh \delta + 2 \coth \delta + \sinh \delta)\}^{\frac{4p \log 4p}{\chi}}, \tag{77}$$

$$B = 8\rho^{-1} \sinh \delta \sinh \rho e^\chi(\cosh \delta + 2 \coth \delta + \sinh \delta), \tag{78}$$

$$C = \frac{11}{2}\delta A, \tag{79}$$

$$\omega = \frac{4p \log 4p}{\chi}, \tag{80}$$

χ being defined in (16).

[22] That the first factor in the upper bound given in (75) is positive may be shown on the basis of the assumption (67) for ϵ.

Since each element of M is contained in at least one of the sets S_M^*, these results enable us to state the following theorem concerning the phase curve g_M generated in M by the above directed H-line segment g_Φ.

THEOREM. *If ϵ denotes a positive number selected in accordance with* (67) *and* (68), *the phase curve g_M comes within a distance ϵ of every element of M. The H-length of g_Φ and therefore that of g_φ can be taken less than λ_q, where λ_q is subject to the inequality* (76). *The right-hand member in* (76) *is then an upper bound for the ergodic function $T(\epsilon)$.*

Remark. The constants A, B, C, ω in (76) all tend to $+\infty$ for $p \to +\infty$, the constant ω, in particular, being bounded by the inequalities

$$\frac{4p \log 4p}{\log (7 + 4\sqrt{3})} < \omega \leqq \frac{4p \log 4p}{\log[(1 + \sqrt{2})(2 + \sqrt{2\sqrt{2} + 1})]}, \tag{81}$$

as may readily be verified from the definitions of ω and χ.

UNIVERSITY OF MARYLAND.

NOTE ON THE SIMULTANEOUS ORTHOGONALITY OF HARMONIC POLYNOMIALS ON SEVERAL CURVES

By J. L. Walsh and G. M. Merriman

1. In the plane of the complex variable $z = x + iy$, the polynomials $1, z, z^2, \cdots$ are mutually orthogonal, not merely on the circumference $|z| = 1$, but also on every circumference $|z| = R$, in the sense that

$$\int_{|z|=R} z^k \bar{z}^l \, |dz| = 0 \qquad k \neq l.$$

The general problem of the existence of sets of polynomials in z which are simultaneously orthogonal, with respect to suitable norm functions, on each of several curves in the z-plane has been studied only recently. Let us say that the set $p_k(z)$ of polynomials in z is *canonical* on a rectifiable Jordan curve C with respect to the norm function $n(z)$ provided the set $p_k(z)$ is found by orthogonalization on C of the set $1, z, z^2, \cdots$ with respect to the positive continuous norm function $n(z)$, and provided the coefficient of z^k in $p_k(z)$ is chosen positive. Walsh established[1] the orthogonality with respect to a suitable norm function of certain Tchebycheff polynomials on *all* ellipses of a given confocal family. Szegö[2] and Walsh[3] showed independently and by widely different methods the fact that if the same set of polynomials $p_k(z)$ is canonical on two distinct curves C and C', then either C' is a curve C_R or C is a curve C'_R;[4] Szegö requires analyticity of C and C'. [Let C be an arbitrary Jordan curve in the z-plane, and let the function $z = \psi(w)$ map the exterior of C onto the exterior of the unit circle $|w| = 1$ in the w-plane so that the points at infinity in the two planes correspond to each other. We denote generically by C_R the image (Kreisbild) in the z-plane of the circle $|w| = R > 1$ under this transformation.] Moreover, Szegö[5] exhibited all sets of polynomials in z, each set canonical simultaneously on *all* C_R of a given family, $1 < R < \infty$.[6] The general problem of the existence of sets of polynomials canonical simultaneously on only two curves

Received October 20, 1936; presented to the American Mathematical Society, December 1936.

[1] Bull. Am. Math. Soc., vol. 40 (1934), pp. 84–88. Also *Interpolation and Approximation*, New York, 1935, p. 134, Theorem 12.

[2] Trans. Am. Math. Soc., vol. 37 (1935), pp. 196–206.

[3] *Interpolation and Approximation*, p. 134, Theorem 11.

[4] The analogous result for harmonic polynomials follows directly by the methods of Walsh (loc. cit. and Trans. Am. Math. Soc., vol. 33 (1931), pp. 370–388, especially p. 385).

[5] Loc. cit.

[6] These sets are enumerated in §2, below.

has not been solved; the second-named writer has, however, some as yet unpublished results on this problem.[7]

The entire theory of expansions in harmonic polynomials[8] in x and y is analogous to, but by no means identical with, the theory of expansions in polynomials in z.[9] It is thus in order to study the general problem of sets of harmonic polynomials orthogonal, with respect to suitable norm functions, simultaneously on several rectifiable Jordan curves; and it is the object of the present note to establish the more immediate results concerning this problem.

We investigate the sets of harmonic polynomials which are obtained by separating into real and pure imaginary parts the sets of polynomials in z which are canonical, with respect to suitable norm functions, on *all* curves C_R of a given family F. Some of these sets of harmonic polynomials are orthogonal, with respect to the same norm functions that were used in connection with the corresponding polynomials in z, simultaneously on *every* C_R of F; the remaining sets are orthogonal on *no* C_R.

Besides this surprising and varied deviation of the facts concerning the orthogonality of these harmonic polynomials from the facts concerning the orthogonality of their generating polynomials in z, there is a further interesting consequence of our discussion: a new property of the orthogonal polynomials in z is derived, which is some cases yields a new formula for the polynomial expansion of an analytic function.

2. We list for reference the sets of polynomials in z which are known to be simultaneously canonical, with respect to suitable norm functions, on all curves C_R of the given families, $1 < R < \infty$; from these sets we derive the sets of harmonic polynomials to be studied. Where the basic curve C is not a circle, we enumerate also the transforms of the polynomials and norm functions under the mapping function $z = \psi(w)$ defined in §1; we denote by Γ_R the circle $|w| = R$, the image of C_R. The real and positive norm function can be expressed in the form $|D(z)|^2 \equiv |\Delta(w)|^2$, $z = \psi(w)$, where $D(z)$ $[\equiv \Delta(w)]$ is known to be analytic and non-vanishing in the extended z- [or w-] plane outside C [the unit circle $|w| = 1$].[10] In all cases the norm function on a curve C_R may be altered by multiplication by a positive constant, and the entire configuration may be subjected to an arbitrary linear integral transformation. The orthogonality condition is chosen in the form

$$\int_{C_R} p_k(z)\overline{p_l(z)}\,|D(z)|^2|dz| = \int_{\Gamma_R} p_k(z)\overline{p_l(z)}\,n(w)\,|dw| = 0 \qquad k \neq l,\ z = \psi(w),$$

[7] Presented to the American Mathematical Society, September, 1936; see Bull. Am. Math. Soc., vol. 42 (1936), Abstract 279.

[8] For a report on this theory, see for instance Walsh, Bull. Am. Math. Soc., vol. 35 (1929), pp. 499–544.

[9] For example, an arbitrary function continuous on a circumference C can be uniformly expanded on C in harmonic polynomials; the analogous proposition for polynomials in z is false.

[10] Szegö, loc. cit., Theorem 1.

where $n(w) \equiv |\Delta(w)|^2 |\psi'(w)|$ serves as the norm function in the w-plane.

The list (Szegö, loc. cit.) follows:

I. The set C_R is the set of concentric circles $|z| = R > 0$;
$D(z) = 1$; $p_k(z) = z^k$.

II.[11] The set C_R is the set of concentric circles, $|z| = R > 1$;
$D(z) = (1 - z^{-\alpha})^{-1}$, α a positive integer;
$$|D(z)|^2 = R^{2\alpha} z^{\alpha}/(z^{\alpha} - 1)(R^{2\alpha} - z^{\alpha});$$
$p_k(z) = z^k$ for $0 \leqq k < \alpha$; $p_k(z) = z^{k-\alpha}(z^{\alpha} - 1)$ for $k \geqq \alpha$.

III.[12] The set C_R is the set of confocal ellipses, foci ± 1;
$D(z) = \{z + (z^2 - 1)^{\frac{1}{2}}\}^{\frac{1}{2}} (z^2 - 1)^{-\frac{1}{4}}$; $z = \psi(w) = \frac{1}{2}(w + w^{-1})$;
$\Delta(w) = \{\frac{1}{2}(1 - w^{-2})\}^{-\frac{1}{2}}$; $n(w) = 1$; $p_k(z) = w^k + w^{-k}$.

IV. The set C_R and $z = \psi(w)$ as in III; $D(z) = \{z + (z^2 - 1)^{\frac{1}{2}}\}^{-\frac{1}{2}}(z^2 - 1)^{\frac{1}{4}}$;
$\Delta(w) = \{\frac{1}{2}(1 - w^{-2})\}^{\frac{1}{2}}$; $n(w) = (w^2 - 1)(R^4 - w^2)/4R^4 w^2$;
$p_k(z) = (w^{k+1} - w^{-k-1})/(w - w^{-1})$.

V. The set C_R and $z = \psi(w)$ as in III; $D(z) = (z - 1)^{\frac{1}{4}}(z + 1)^{-\frac{1}{4}}$;
$\Delta(w) = (1 - w^{-1})^{\frac{1}{2}}(1 + w^{-1})^{-\frac{1}{2}}$; $n(w) = (w - 1)(R^2 - w)/2R^2 w$;
$p_k(z) = (w^{k+\frac{1}{2}} - w^{-k-\frac{1}{2}})/(w^{\frac{1}{2}} - w^{-\frac{1}{2}})$.

We shall suffix the subscript H (e.g., $\mathrm{I_H}$) to the above numerals I–V to designate the harmonic polynomials derived from the sets I–V by separation into real and pure imaginary parts. We reiterate that in the cases III, IV, V the notation $n(w)$ does not represent merely the transform of the norm function $n(z) \equiv |D(z)|^2$ in the z-plane, but represents $n(z)\,|\psi'(w)| \equiv |\Delta(w)|^2 |\psi'(w)|$, the norm function in the w-plane.

3. It would perhaps seem most natural to offer real-variable proofs of our results concerning the harmonic polynomials $\mathrm{I_H}$–$\mathrm{V_H}$, using trigonometric forms of both polynomials and norm functions. Although such proofs exist, it seems simpler and more fruitful to carry out the discussion by methods of complex variables. For this purpose we first proceed to obtain some general theorems concerning orthogonality of polynomials in z and of the harmonic polynomials which result from separating them into real and pure imaginary parts.

Let the set of polynomials Σ: $\{p_k(z)\}$, canonical or not, be orthogonal with respect to the positive continuous norm function $n(z)$, on a given rectifiable Jordan curve C:

$$\int_C p_k(z)\overline{p_l(z)}\, n(z)\,|dz| = 0 \qquad k \neq l. \tag{1}$$

If each polynomial $p_k(z)$ is separated into real and pure imaginary parts, $p_k \equiv p_k' + ip_k''$, we obtain a set of polynomials in x and y each of which is harmonic in the entire plane:

[11] This set was exhibited by Szegö (Math. Ann., vol. 79 (1919), pp. 323–339), but without mention at that time of orthogonality on more than one curve. Compare Walsh, Mémorial des sciences math., Fasc. 73, p. 43.

[12] This is the Tchebycheff set proved by Walsh to be orthogonal simultaneously on all C_R.

$$S: \left\{ \begin{matrix} p_0', p_1', p_2', \cdots \\ p_0'', p_1'', p_2'', \cdots \end{matrix} \right\}.$$

If $p_0(z)$ is real and of degree zero, which necessarily occurs if the set Σ is canonical, the polynomial p_0'' vanishes identically and may be omitted here. An *arbitrary* set of form S is orthogonal on C with respect to the norm function $n(z)$, provided that the three sets of conditions

$$\int_C p_k' p_l' n(z) \mid dz \mid = 0 \qquad k \neq l,$$

$$\int_C p_k'' p_l'' n(z) \mid dz \mid = 0 \qquad k \neq l, \tag{2}$$

$$\int_C p_k' p_l'' n(z) \mid dz \mid = 0,$$

are satisfied. Under the present assumption of the orthogonality of the set Σ, the second conditions in (2) are satisfied when and only when the first conditions are satisfied; thus, for harmonic polynomials obtained by separation of orthogonal polynomials in z into real and pure imaginary parts, the orthogonality conditions (2) are equivalent to

$$\int_C p_k' p_l' n(z) \mid dz \mid = 0 \text{ for } k \neq l, \qquad \int_C p_k' p_l'' n(z) \mid dz \mid = 0. \tag{3}$$

We observe that if conditions (1) and (3) are satisfied, there are also satisfied the conditions

$$\int_C p_k(z) p_l(z) n(z) \mid dz \mid = 0 \qquad k \neq l. \tag{4}$$

Conversely, if both (1) and (4) are satisfied, and if in addition

$$\int_C p_k' p_l'' \, n(z) \mid dz \mid = 0 \qquad k = l, \tag{5}$$

we may infer the satisfaction of conditions (3). Conditions (5) may be replaced by a condition on the pure imaginary component of an integral:

$$\Im \left\{ \int_C [p_k(z)]^2 n(z) \mid dz \mid \right\} = 0 \qquad k = 0, 1, 2, \cdots. \tag{6}$$

We may thus state

THEOREM 1. *If the set* $\{p_k\}$ *of polynomials in* z *is orthogonal on a rectifiable curve* C *with respect to the positive continuous norm function* $n(z)$, *then it is necessary and sufficient for the orthogonality on* C *with respect to* $n(z)$ *of the set of harmonic polynomials* (p_k', p_k'') *obtained by separating each* $p_k(z)$ *into real and pure imaginary parts that conditions* (4) *and* (6) [*or* (5)] *be satisfied; that is to say, it is necessary and sufficient that the integral*

$$J = \int_C p_k(z) p_l(z) n(z) \, | dz |$$

vanish for $k \neq l$ *and be real for* $k = l$.

It is easy to prove, by combining formally conditions (2) so as to yield conditions (1),

THEOREM 2. *If the set* $\{p_k', p_k''\}$ *of harmonic polynomials, where* p_k' *and* p_k'' *are conjugate harmonic functions, is orthogonal on a rectifiable curve* C *with respect to a positive continuous norm function* $n(z)$, *then the set* $\{p_k \equiv p_k' + ip_k''\}$ *of polynomials in* z *is also orthogonal on* C *with respect to* $n(z)$.

It is Theorem 1 to which we shall appeal in our study of the orthogonality properties of the harmonic polynomials I_n–V_n. The integral J, however, is also of interest and importance in connection with an expansion problem. Let the function $f(z)$ be analytic on and within C; the formal expansion of $f(z)$ on C in terms of the polynomials $\{p_k(z)\}$, now assumed canonical on C, is

$$f(z) \sim \sum_{j=0}^{\infty} a_j p_j(z), \tag{7}$$

where

$$a_j = \int_C f(z)\, \overline{p_j(z)}\, n(z) \, | dz | \Big/ \int_C p_j(z)\, \overline{p_j(z)}\, n(z) \, | dz |, \quad j = 0, 1, 2, \cdots ; \tag{8}$$

of course we have the relation

$$\int_C p_j(z)\, \overline{p_j(z)}\, n(z) \, | dz | \neq 0 \qquad j = 0, 1, 2, \cdots . \tag{9}$$

Under the present hypothesis, the relation (7) is an actual *equation*, valid uniformly on and within C.[13]

If now conditions (4) are also satisfied by the set $\{p_k(z)\}$, we can obtain an alternative expression for the coefficients a_j valid under certain circumstances. Multiplication of *equation* (7) by $p_j(z)n(z)$ and integration term by term over C yields

$$a_j = \int_C f(z)\, p_j(z)\, n(z) \, | dz | \Big/ \int_C [p_j(z)]^2\, n(z) \, | dz | \qquad j = 0, 1, 2, \cdots , \tag{10}$$

provided of course that

$$\int_C [p_j(z)]^2 n(z) \, | dz | \neq 0 \qquad j = 0, 1, 2, \cdots . \tag{11}$$

We collect these results as

[13] This theorem is due to Szegö in the case $n(z) \equiv 1$ with C analytic, to Smirnoff in the case $n(z) \equiv 1$ with C rectifiable and satisfying an auxiliary condition, and to Walsh in the general case that both $n(z)$ and C are arbitrary. See Walsh, *Interpolation and Approximation*, §5.2.

THEOREM 3. *If the set of polynomials $\{p_k(z)\}$ is canonical on C with respect to a norm function $n(z)$, and also satisfies conditions* (4) *and* (11) (*that is to say, if the integral J vanishes for $k \neq l$ but not for $k = l$*), *then an arbitrary function $f(z)$ analytic on and within C has the formal expansion* (7), *where a_j is given both by* (8) *and by* (10). *The expansion* (7) *is valid uniformly on and within C.*

4. The general results established in §3 will now be applied to the study of the orthogonality with respect to the given norm functions, of the sets I_H–V_H of harmonic polynomials, on all C_R of the given families of curves. The computation will be carried out in the plane in which the family of curves is a set of concentric circles, that is, after the usual mapping $z = \psi(w)$ in cases III_H–V_H. Let Γ_R designate the circle $|w| = R > 1$; Γ_R coincides with C_R in cases I_H and II_H and is its image in cases III_H–V_H.

In case I_H,

$$J = \int_{\Gamma_R} z^{k+l}\,|dz| = \frac{R}{i}\int_{\Gamma_R} z^{k+l-1}dz.$$

Hence $J = 0$ for all choices of k and l except $k = l = 0$; in the latter case, however, equation (5) is satisfied, since p_0 is real. Thus conditions (4) and (6) of Theorem 1 are completely satisfied, and we have established the well-known fact that the harmonic polynomials are orthogonal simultaneously on all C_R. But condition (11) is not satisfied, so formulas (10) for the coefficients a_j are not valid.

In case III_H (the study of II_H is more complex and we postpone it temporarily),

$$\begin{aligned} J &= \int_{\Gamma_R} (w^k + w^{-k})(w^l + w^{-l})\,|dw| \\ &= \frac{R}{i}\int_{\Gamma_R} [w^{k+l-1} + w^{k-l-1} + w^{-k+l-1}\, w^{-k-l-1}]\,dw. \end{aligned}$$

Hence $J = 0$ immediately if $k \neq l$, so conditions (4) are satisfied. If $k = l$, we have

$$J = \frac{cR}{i}\int_{\Gamma_R} \frac{dw}{w} = 2\pi cR,$$

where $c = 4$ if $k = l = 0$; otherwise $c = 2$. Since this value of J is real and non-vanishing, conditions (6) and (11) are satisfied. We conclude from Theorem 1 that *the set* III_H *is orthogonal simultaneously on every* C_R, and from Theorem 3 that *formulas* (10) *for the coefficients a_j in* (7) *are valid.*[14]

[14] A form of orthogonality other than (1) and (4) is also of interest here, namely

$$\int_C Q(z)\; p_k(z)\; p_l(z)\; dz = 0 \qquad\qquad k \neq l,$$

where $Q(z)$ is suitably chosen. See Geronimus, Trans. Amer. Math. Soc., vol. 33 (1931), pp. 322–328.

5. Among the sets ($\alpha = 1, 2, \cdots$) of polynomials Π_H, the sets Π_H, $\alpha > 2$ are orthogonal on *no* C_R; the sets Π_H, $\alpha = 1$ or 2 are orthogonal on *every* C_R.

We have the following three expressions for J for the three possible choices of k and l relative to α:

$$J_1 = \frac{R^{2\alpha+1}}{i} \int_{\Gamma_R} \frac{z^{k+l-\alpha-1}(z^\alpha - 1)\,dz}{R^{2\alpha} - z^\alpha} \qquad k \geqq \alpha,\ l \geqq \alpha;$$

$$J_2 = \frac{R^{2\alpha+1}}{i} \int_{\Gamma_R} \frac{z^{k+l-1}\,dz}{R^{2\alpha} - z^\alpha} \qquad k \geqq \alpha,\ l < \alpha;$$

$$J_3 = \frac{R^{2\alpha+1}}{i} \int_{\Gamma_R} \frac{z^{k+l+\alpha-1}\,dz}{(z^\alpha - 1)(R^{2\alpha} - z^\alpha)} \qquad k < \alpha,\ l < \alpha.$$

The integrals J_1 and J_2 both vanish, for all pertinent choices of k and l, by Cauchy's integral theorem, since there are no singularities of the integrands interior to Γ_R. The integral J_3, however, possesses singularities at points inside Γ_R, namely at the α-th roots of unity; we evaluate J_3 in the following manner.

Lemma 1. *Let ω be a primitive α-th root of unity. The sum $\sum_{m=1}^{\alpha} \omega^{mq}$ has the value α or 0 according as q is or is not a multiple of α.*

The proof is left to the reader.

The value of the integral J_3, aside from the constant factor, is the sum of the residues of the integrand at the points $\omega, \omega^2, \cdots, \omega^\alpha = 1$. At the point ω^m the residue is

$$\begin{aligned} 2\pi i &\left[\frac{z^{k+l+\alpha-1}}{R^{2\alpha} - z^\alpha} \cdot \frac{z - \omega^m}{z^\alpha - 1}\right]_{z=\omega^m} \\ &= 2\pi i \left[\frac{z^{k+l+\alpha-1} + (k + l + \alpha - 1)(z - \omega^m) z^{k+l+\alpha-2}}{\alpha(R^{2\alpha} - z^\alpha) z^{\alpha-1} - \alpha(z^\alpha - 1) z^{\alpha-1}}\right]_{z=\omega^m} \\ &= 2\pi i \frac{\omega^{m(k+l)}}{\alpha(R^{2\alpha} - 1)}. \end{aligned}$$

Hence the integral J_3 has the value

$$J_3 = \frac{2\pi R^{2\alpha+1}}{\alpha(R^{2\alpha} - 1)} \sum_{m=1}^{\alpha} \omega^{m(k+l)} \tag{12}$$

$$= \begin{cases} 0, \text{ if } k + l \text{ is not a multiple of } \alpha, \\ 2\pi R^{2\alpha+1}/(R^{2\alpha} - 1), \text{ if } k + l \text{ is a multiple of } \alpha. \end{cases}$$

Concerning the polynomials Π_H, $\alpha = 1$, we conclude immediately that the conditions of Theorem 1 are completely satisfied: $J_1 = 0$, $J_2 = 0$ in all circumstances; the integral J_3 is of significance only when $k = l = 0$, in which case the condition on J_3 is (5), fulfilled by virtue of the reality of p_0. *These polynomials Π_H, $\alpha = 1$, are therefore orthogonal simultaneously on every C_R.* On

the other hand, conditions (11) fail for $j > \alpha$, so Theorem 3 is of no significance here.

We turn to the set II_H, $\alpha = 2$. By (12) we have $J_3 = 0$ for $k = 1$, $l = 0$; and J_3 is real for $k = l = 0$ or 1. These are the only cases not covered by $J_1 = J_2 = 0$; thus the conditions of Theorem 1 are satisfied, and *the polynomials* II_H, $\alpha = 2$, *are orthogonal simultaneously on every* C_R'. Again Theorem 3 fails to be applicable.

But if we examine any set II_H, $\alpha > 2$, there always exist pairs of polynomials of the set which *fail to satisfy conditions* (4). For example if we choose $k = \alpha - 1$, $l = 1$, we find from (12) that $J_3 \neq 0$, $k \neq l$; hence conditions (4) of Theorem 1 fail to be satisfied, so *the polynomials* II_H, $\alpha > 2$ *are mutually orthogonal on no* C_R.

6. We turn now to the proof of the orthogonality of the set IV_H on all C_R. We have

$$J = \frac{1}{4R^3 i}\int_{\Gamma_R} \frac{w^{k+1} - w^{-k-1}}{w - w^{-1}}\,\frac{w^{l+1} - w^{-l-1}}{w - w^{-1}}\,\frac{(w^2 - 1)(R^4 - w^2)\,dw}{w^3}$$
$$= \frac{1}{4R^3 i}\int_{\Gamma_R} \frac{(w^{2k} + w^{2k-2} + w^{2k-4} + \cdots + 1)(w^{2l+2} - 1)(R^4 - w^2)\,dw}{w^{k+l+3}}.$$

We expand the numerator, and evaluate the corresponding integrals separately. Suppose first $k \neq l$; it is no loss of generality to assume, as we do, $l > k$. The terms of J resulting from

$$(w^{2k} + w^{2k-2} + \cdots + 1)w^{2l+2}(R^4 - w^2)$$

vanish because in this product each exponent of w is greater than $k + l + 2$; we have $2l + 2 > k + l + 2$. The terms of J resulting from

$$(w^{2k} + w^{2k-2} + \cdots + 1)(R^4 - w^2)$$

vanish because in this product each exponent of w is less than $k + l + 2$; we have $k + l + 2 > 2k + 2$. Hence conditions (4) are satisfied. Suppose next $k = l$; the only terms in J which do not obviously vanish are those contributed by the terms $R^4 w^{2l+2}$ and w^{2k+2} in the numerator of the integrand. We have

$$J = \frac{\pi(R^4 + 1)}{2R^3}.$$

Hence conditions (6) of Theorem 1 are satisfied, so *the set* IV_H *is orthogonal simultaneously on every* C_R. Conditions (11) of Theorem 3 are also satisfied, so that theorem is applicable.

Methods similar to these serve to prove that the set V_H is orthogonal simultaneously on every C_R, and that Theorem 3 is applicable; we omit further details.

7. One further pertinent deduction is to be made, from Theorem 2 of §3. In view of our study of the orthogonality, on all C_R, of the sets of harmonic

polynomials I_H–V_H, and in view of the fact that the polynomials I–V are the only sets of polynomials in z which are canonical simultaneously on all C_R of a given family, Theorem 2 yields the conclusion that the sets I_H, II_H for $\alpha = 1$ or 2, III_H–V_H are the *only* sets of harmonic polynomials orthogonal simultaneously on every C_R of a given family which can be obtained by separation into real and pure imaginary parts of canonical sets of polynomials in z.

The case of the polynomials II_H, $\alpha > 2$, illustrates the falsity of the converse of Theorem 2.

8. We remark that the results just proved on the orthogonality of the sets I_H–V_H can also be established by the use of real variables. Various integrals that present themselves are analogous to Poisson's integral, and can be evaluated at once by identifying the given integrals with Poisson's integral for suitably chosen integrands and suitably chosen values of the parameters; in each case the corresponding Dirichlet problem has for its solution a fairly obvious harmonic polynomial. It might be supposed that the real-variable proofs are preferable and more natural in the study of harmonic functions, but as a matter of fact the methods of the complex variable as we have used them are both more fruitful (see for example our Theorem 3) and more illuminating as to the general structure of the harmonic polynomials studied (see conditions (4) and (6) of Theorem 1).

For the sake of reference we list, in trigonometric form, the harmonic polynomials I_H–V_H, together with their respective norm functions:

I_H. $p_k'(z) = R^k \cos k\theta,\ p_k''(z) = R^k \sin k\theta;\ n(z) = 1;$

$$z = R(\cos\theta + i \sin\theta).$$

II_H. $p_k'(z) = R^k \cos k\theta,\ p_k''(z) = R^k \sin k\theta$, for $0 \leqq k < \alpha$,

α a positive integer;

$$p_k'(z) = R^k \cos k\theta - R^{k-\alpha} \cos (k-\alpha)\theta,$$

$$p_k''(z) = R^k \sin k\theta - R^{k-\alpha} \sin (k-\alpha)\theta, \text{ for } k \geqq \alpha;$$

$$n(z) = R^{2\alpha}/(1 - 2R^\alpha \cos \alpha\theta + R^{2\alpha});\ z = R(\cos\theta + i\sin\theta).$$

III_H. $p_k'(z) = (R^k + R^{-k}) \cos k\theta,\ p_k''(z) = (R^k - R^{-k}) \sin k\theta;$

$$n(w) = 1;\ w = R(\cos\theta + i\sin\theta);\ z = \psi(w) = \tfrac{1}{2}(w + w^{-1}).$$

IV_H. $p_k'(z) = R^2\{(R^{k+1} - R^{-k-1})(R - R^{-1}) \cos (k+1)\theta \cos\theta$

$$+ (R^{k+1} + R^{-k-1})(R + R^{-1}) \sin (k+1)\theta \sin\theta\}/(1 - 2R^2 \cos 2\theta + R^4);$$

$$p_k''(z) = R^2\{(R^{k+1} + R^{-k-1})(R - R^{-1}) \sin (k+1)\theta \cos\theta$$

$$- (R^{k+1} - R^{-k-1})(R + R^{-1}) \cos (k+1)\theta \sin\theta\}/(1 - 2R^2 \cos 2\theta + R^4);$$

$$n(w) = (1 - 2R^2 \cos 2\theta + R^4)/4R^4;\ w = R(\cos\theta + i\sin\theta).$$

$$\mathrm{V}_{\mathrm{II}}.\quad p_k'(z) = R\{(R^{k+\frac12} - R^{-k-\frac12})(R^{\frac12} - R^{-\frac12}) \cos (k + \tfrac12)\theta \cos \tfrac12\theta$$
$$+ (R^{k+\frac12} + R^{-k-\frac12})(R^{\frac12} + R^{-\frac12}) \sin (k + \tfrac12)\theta \sin \tfrac12\theta\}/(1 - 2R \cos \theta + R^2);$$
$$p_k''(z) = R\{(R^{k+\frac12} + R^{-k-\frac12})(R^{\frac12} - R^{-\frac12}) \sin (k + \tfrac12)\theta \cos \tfrac12\theta$$
$$- (R^{k+\frac12} - R^{-k-\frac12})(R^{\frac12} + R^{-\frac12}) \cos (k + \tfrac12)\theta \sin \tfrac12\theta\}/(1 - 2R \cos \theta + R^2);$$
$$n(w) = (1 - 2R \cos \theta + R^2)/2R^2;\ w = R(\cos \theta + i \sin \theta).$$

HARVARD UNIVERSITY AND THE UNIVERSITY OF CINCINNATI.

CERTAIN TERNARY CUBIC ARITHMETICAL FORMS

By E. T. Bell

1. By an obvious oversight, G. B. Mathews stated (without restriction on the integer n) that if the integer m is representable by integers x, y, z in the form

$$x^3 + ny^3 + n^2z^3 - 3nxyz,$$

it can be represented in an infinity of ways.[1] If n is the cube of a rational integer, and $m \neq 0$, the number of representations, if any, is finite. We shall show how these representations may be found. In certain simple cases (§5) it is possible to find the exact number of representations; in all cases we give an upper bound (§4) for this number.

2. We consider the representations by integers x, y, z of the integer m in the form

$$x^3 + t^3y^3 + t^6z^3 - 3t^3xyz, \tag{1}$$

where t is a constant integer $\neq 0$. If $m = 0$, we have the infinity of representations $(x, y, z) = (ht^2, ht, h)$, where h is an arbitrary integer (and possibly further representations).

Henceforth we shall take $m \neq 0$. Let $m = d\delta$, where d, δ are integers. From (1) we have

$$(x + ty + t^2z)(x^2 + t^2y^2 + t^4z^2 - txy - t^2xz - t^3yz) = d\delta;$$

hence we may take

$$x + ty + t^2z = d, \tag{2}$$

and equate the second factor to δ. Replacing t^2z by $d - x - ty$, we get

$$3(x^2 + txy + t^2y^2 - dx - dty) + d^2 - \delta = 0.$$

Thus

$$d^2 - \delta \equiv 0 \bmod 3; \qquad d^2 - \delta = 3h, \tag{3}$$

where h is an integer, and

$$x^2 + (ty - d)x + t^2y^2 - dty + h = 0. \tag{4}$$

Received November 19, 1936.

[1] G. B. Mathews, Proceedings of the London Mathematical Society, vol. 21 (1891), pp. 280–7; see also Dickson's *History of the Theory of Numbers*, vol. 2, 1920, p. 594. In applying Dirichlet's theory of representations by norms of algebraic integers, Mathews omitted to state, loc. cit., p. 281 (as he evidently intended) that the real cube root of the n he is considering must be irrational.

In order that (4) shall have an integer root x it is necessary that the discriminant be an integer square a^2; whence

(5) $$3t^2y^2 - 2dty + a^2 + 4h - d^2 = 0;$$

and in order that (5) shall have an integer root y it is necessary that the discriminant be an even integer square $4b_1^2$; thus

$$d^2t^2 - 3t^2(a^2 + 4h - d^2) = b_1^2.$$

It follows that $t \mid b_1$, say $b_1 = tb$. Hence, by (3),

$$3a^2 + b^2 = 4\delta;$$

and from (4), (5), (2) we find the values of x, y, z stated in the following

THEOREM 1. *All integer solutions* (x, y, z), *if any, of*

$$x^3 + t^3y^3 + t^6z^3 - 3t^3xyz = m$$

in which t is a constant integer $\neq 0$ *and m is any integer* $\neq 0$, *are given by*

(6) $$6x = 2d - b + 3a, \qquad 3ty = d + b, \qquad 6t^2z = 2d - b - 3a,$$

where a, b are integers such that

(7) $$3a^2 + b^2 = 4\delta,$$

and d, δ are integers such that

(8) $$m = d\delta, \qquad d^2 \equiv \delta \qquad \text{mod } 3.$$

Hence m, d are both positive or both negative, and

(9) $$(d, \delta) \equiv (0, 0), (1, 1), (-1, 1) \qquad \text{mod } 3.$$

THEOREM 2. *The only* $m \neq 0$ *representable in the form* (1) *are of the forms* $9n$, $3n \pm 1$.

3. Taking the first of (9), we have $(d, \delta) = (3d_1, 3\delta_1)$. Hence, from the first of (6), $3 \mid b$, $b = 3b_1$;

$$x = d_1 - \tfrac{1}{2}(b_1 - a), \qquad b_1 \equiv a \qquad \text{mod } 2,$$

the last since x is an integer. From (7) we now have $a^2 + 3b_1^2 = 4\delta_1$. Taking

$$(a, b_1) = (2a_2, 2b_2), (2a_3 + 1, 2b_3 + 1),$$

and dropping all suffixes, we get from Theorem 1,

THEOREM 3. *All integer solutions* (x, y, z), *if any, of*

$$x^3 + t^3y^3 + t^6z^3 - 3t^3xyz = 9n, \qquad n \neq 0,$$

are given by

$$x = d - b + a, \qquad ty = d + 2b, \qquad t^2z = d - b - a,$$

with a, b, d, δ integers such that

$$n = d\delta, \qquad a^2 + 3b^2 = \delta, \tag{10}$$

and

$$x = d - b + a, \qquad ty = d + 2b + 1, \qquad t^2z = d - b - a - 1,$$

with a, b, d, δ integers such that

$$n = d\delta, \qquad (2a + 1)^2 + 3(2b + 1)^2 = 4\delta. \tag{11}$$

Similarly, the second and third of (9) give the following

THEOREM 4. *All integer solutions (x, y, z), if any, of*

$$x^3 + t^3y^3 + t^6z^3 - 3t^3xyz = n, \qquad n \equiv 1 \mod 3,$$

are given by

$$x = (d - 1)/3 - b + a + 1, \qquad ty = (d - 1)/3 + 2b,$$

$$t^2z = (d - 1)/3 - b - a,$$

with a, b, d, δ integers such that

$$n = d\delta, \qquad d \equiv \delta \equiv 1 \mod 3, \tag{12}$$

$$3(2a + 1)^2 + (6b - 1)^2 = 4\delta.$$

THEOREM 5. *All integer solutions (x, y, z), if any, of*

$$x^3 + t^3y^3 + t^6z^3 - 3t^3xyz = n, \qquad n \equiv -1 \mod 3,$$

are given by

$$x = (d + 1)/3 - b + a, \quad ty = (d + 1)/3 + 2b, \quad t^2z = (d + 1)/3 - b - a - 1,$$

with a, b, d, δ integers such that

$$n = d\delta, \qquad d \equiv -\delta \equiv -1 \mod 3, \tag{13}$$

$$3(2a + 1)^2 + (6b + 1)^2 = 4\delta.$$

4. An upper bound to the number $F(n)$ of representations (x, y, z) of n in the form (1) is easily obtained from the number $N(n)$ of representations (a, b) of n in the form $a^2 + 3b^2$. Let $n = 2^\alpha m$, where m is odd. Then[2] if $E(m)$ denotes the excess of the number of divisors $3h + 1$ of m over the number of divisors $3h - 1$,

$$N(n) = 0, \qquad \text{if } \alpha \text{ is odd};$$
$$N(n) = 2E(m), \qquad \text{if } \alpha = 0;$$
$$N(n) = 6E(m), \qquad \text{if } \alpha \text{ is even}, > 0.$$

Referring to (10), (11), we have $(n \neq 0)$,

$$F(9n) \leqq 6\left[\sum N(\delta) + \sum N(4\delta)\right],$$

[2] L. E. Dickson, *Introduction to the Theory of Numbers*, 1929, p. 80, Exercise 3.

where the sums refer to all divisors $\delta > 0$ of $2^{\alpha}m$. These divisors are $2^{\beta}\delta'$, where $0 \leqq \beta \leqq \alpha$, and δ' is any divisor > 0 of m. A short reduction of the resulting inequality for $F(9n)$ gives

THEOREM 6. *If $F(k)$ denotes the total number of representations of k in the form* (1),

$$F(9n) \leqq [6\alpha + 5 + 3(-1)^{\alpha}] \sum E(\delta),$$

where $n = 2^{\alpha}m$, $\alpha \geqq 0$, m is odd, and $\sum$ refers to the divisors $\delta > 0$ of m.

The factor 6 enters, since a representation (x, y, z) for a particular (a, b) contributes at most 6 representations by permutations of x, y, z. In the same way, (12), (13) give

THEOREM 7. *If* $n \equiv \pm 1 \bmod 3$,

$$F(n) \leqq 12 \sum E(\delta),$$

where $\sum$ refers to all divisors $\delta > 0$ of n.

5. Taking $t = 1$ in (1) we now consider representations in

$$x^3 + y^3 + z^3 - 3xyz. \tag{14}$$

THEOREM 8. *All integers* $\equiv 0 \bmod 9$, *or* $\equiv \pm 1 \bmod 3$, *and only these, are represented in the form* (14), *and for integers* $\neq 0$ *the number of representations is finite.*

This follows from Theorem 2 and (10), (11), (12), (13), since each of these equations has at least one integer solution when $\delta = 1$. The first part of this theorem was proved otherwise by Carmichael,[3] who showed also that for integers > 0 in each case there is a representation (x, y, z) with x, y, z non-negative. Here, taking $t = 1$ in Theorems 3, 4, 5, we reach the same conclusion, with the additional information in

THEOREM 9. *In the form* (14), $9n$ *has the representation* $(n + 1, n, n - 1)$, *and if* $n > 1$, *a representation* (x, y, z) *with* $x > 0, y > 0, z > 0$; $n \equiv 1 \bmod 3$ *has the representation* $\left(\frac{n+2}{3}, \frac{n-1}{3}, \frac{n-1}{3}\right)$, *and if* $n > 1$, *a representation* (x, y, z) *with* $x > 0$, $y > 0$, $z > 0$; $n \equiv -1 \bmod 3$ *has the representation* $\left(\frac{n+1}{3}, \frac{n+1}{3}, \frac{n-2}{3}\right)$, *and if* $n > 2$, *a representation* (x, y, z) *with* $x > 0$, $y > 0, z > 0$.

Considerably more may be stated for certain special forms of n. With N, E as in §4, let p be a prime > 0. Then $N(p) = 4$ or 0 according as $p \equiv 1 \bmod 3$ or $p \equiv -1 \bmod 3$. Hence for $p \equiv 1 \bmod 3$ there is precisely one representation $(a, b) = (\alpha, \beta)$, with $\alpha > 0, \beta > 0$, of p in the form $a^2 + 3b^2$. A straightforward application of these remarks to Theorems 3, 4, 5 with $t = 1$ gives

THEOREM 10. *If p is a positive prime* $\equiv 1 \bmod 3$, *there are precisely* 18 *repre-*

[3] R. D. Carmichael, Bulletin of the American Mathematical Society, vol. 22 (1915), pp. 111–117.

sentations (x, y, z) *of* $9p$ *in the form* $x^3 + y^3 + z^3 - 3xyz$, *obtained by permutations of* x, y, z *in*

$$\begin{aligned}(x, y, z) = {} & (p - 1, p, p + 1),\\ & (1 + \alpha + \beta, 1 - 2\beta, 1 - \alpha + \beta),\\ & (1 + \alpha - \beta, 1 + 2\beta, 1 - \alpha - \beta),\end{aligned}$$

where (α, β) *is the (necessarily) unique representation* $(a, b) = (\alpha, \beta)$ *of* p *in the form* $a^2 + 3b^2$ *with* $a > 0$, $b > 0$; *if* p *is a positive prime* $\equiv -1 \bmod 3$, *there are precisely* 6 *representations of* $9p$, *obtained by permutations from* $(p - 1, p, p + 1)$.

The implied distinctness of these representations is easily seen from the primality of p and a simple contradiction. The corresponding theorems for $p < 0$ follow at once if we note that δ in preceding theorems is > 0. Similarly we find

THEOREM 11. *The only representations of the positive prime* $p \equiv -1 \bmod 3$ *in the form* $x^3 + y^3 + z^3 - 3xyz$ *are the three obtained by permutations from* $\left(\frac{p+1}{3}, \frac{p+1}{3}, \frac{p-2}{3}\right)$.

The following immediate consequences of Theorem 7 are of some interest.

THEOREM 12. *If* p *is a positive prime,* α *an integer* $\geqq 0$, *and if* $G(n)$ *denotes the total number of representations of* n *in the form* (14),

$$G(p^\alpha) \leqq 6(\alpha + 1)(\alpha + 2), \qquad p \equiv 1 \quad \bmod 3;$$

$$G(p^\alpha) \leqq 3[2\alpha + 3 + (-1)^\alpha], \qquad p \equiv -1 \quad \bmod 3.$$

The first part of Theorem 8 (Carmichael's result) has been generalized by Hua[4] to any circulant.

CALIFORNIA INSTITUTE OF TECHNOLOGY.

[4] L. K. Hua, Tohôku Mathematical Journal, vol. 39 (1934), pp. 316–321.

ELEMENTARY PROOFS OF SOME KNOWN THEOREMS OF THE THEORY OF COMPLEX EUCLIDEAN SPACES

By M. H. Stone and J. D. Tamarkin

In this note we shall indicate strictly elementary proofs of certain well-known fundamental theorems concerning a complex Euclidean space $\mathfrak{L}$.[1] Such proofs are not difficult to construct and must be widely known, but the proofs actually available in the literature repose, in many cases, upon arguments drawn from topology or from the spectral theory. It may therefore be helpful to put on record proofs of a more elementary nature.

We begin with a proposition due, in its general form, to Banach.[2]

Theorem 1. *If the sequence $\{g_n\}$ in $\mathfrak{L}$ has the property that $\{(f, g_n)\}$ is a bounded sequence for each f in $\mathfrak{L}$, then the sequence $\{ | g_n | \}$ is also bounded.*

In constructing a proof by contradiction there is no loss of generality in assuming that $(f, g_n) \to 0$ for each f in $\mathfrak{L}$. Indeed, if we make the assumption that $\{ | g_n | \}$ is not bounded, we can select a subsequence $\{g'_n\}$ of $\{g_n\}$ such that $| g'_n | \geqq n^2$, and we can then consider the sequence $\{g''_n\}$, where $g''_n = g'_n/n$, for which the properties $| g''_n | \geqq n$, $(f, g''_n) \to 0$ are evident.

We therefore proceed to obtain a contradiction from the assumption that $| g_n | \to \infty$, $(f, g_n) \to 0$ for each f in $\mathfrak{L}$. Now by selecting an appropriate subsequence and renumbering its elements, we may further suppose that

$$| g_n | \geqq 2^n, \qquad | (g_m, g_n) | \leqq 1 \qquad \text{when } m \neq n.$$

In fact, if we have chosen $g_1, g_2, \cdots, g_N$, the relations $| g_n | \to \infty$, $(g_m, g_n) \to 0$ for $m = 1, 2, \cdots, N$ allow us to choose g_{N+1} from the remaining members of the original sequence so that

$$| g_{N+1} | \geqq 2^{N+1}, \qquad | (g_m, g_{N+1}) | \leqq 1 \qquad \text{for } m = 1, 2, \cdots, N.$$

An obvious inductive construction therefore provides us with the desired subsequence. The inequality $| g_n / | g_n |^2 | \leqq 2^{-n}$ shows that the series $\sum_{\nu=1}^{\infty} g_\nu / | g_\nu |^2$ is dominated (in norm) term for term by the convergent series $\sum_1^{\infty} 2^{-\nu}$, and therefore is convergent in $\mathfrak{L}$ to a limit f. By assumption $(f, g_n) \to 0$. On the other hand, we have the contradictory inequality which follows:

Received December 30, 1936.

[1] By a complex Euclidean space we mean a complex linear vector space with definite Hermitian-bilinear inner product (f, g), complete in the metric defined by $| f - g | = (f - g, f - g)^{\frac{1}{2}}$. Hilbert space is a special case.

[2] Banach, *Théorie des Opérations Linéaires*, Warsaw, 1932, p. 80.

$$|(f, g_n)| = \left| \sum_{\nu=1}^{\infty} (g_\nu, g_n)/|g_\nu|^2 \right| = \left| 1 + \sum_{\nu \neq n} (g_\nu, g_n)/|g_\nu|^2 \right|$$

$$\geqq 1 - \sum_{\nu \neq n} |(g_\nu, g_n)|/|g_\nu|^2 \geqq 1 - \sum_{\nu=1}^{\infty} 4^{-\nu} = \frac{2}{3}.$$

Consequently, our theorem is established.

As an immediate corollary of this result, we have

THEOREM 2. *If $\{(f, g_n)\}$ converges for each f in $\mathfrak{L}$, then there exists a uniquely determined element g in $\mathfrak{L}$ such that $(f, g_n) \to (f, g)$. This element g belongs to the closed linear manifold $\mathfrak{M}$ generated by the sequence $\{g_n\}$ and satisfies the inequality $|g| \leqq G$, where $G = \liminf |g_n| < +\infty$.*

Theorem 1 shows at once that G is finite. Now the limit $L(f)$ of (f, g_n) obviously depends linearly on f and, by virtue of the Schwarz inequality $|(f, g_n)| \leqq |f|\,|g_n|$, satisfies the inequality $|L(f)| \leqq G|f|$. A theorem of Riesz[3] therefore asserts the existence of a unique element g in $\mathfrak{L}$ such that $L(f) = (f, g)$ for every f in $\mathfrak{L}$, and such, furthermore, that $|g| \leqq G$. Since f is in the orthogonal complement $\mathfrak{M}^*$ of $\mathfrak{M}$ if and only if $(f, g_n) = 0$ for every n, we see that every such f satisfies $(f, g) = L(f) = 0$. Thus g belongs to the orthogonal complement of $\mathfrak{M}^*$—that is, to $\mathfrak{M}^{**} = \mathfrak{M}$.

It should be noted that the inequality $|g| < G$ can hold; for example, if $\{g_n\}$ is an infinite orthonormal set, we have $g = 0$, $G = 1$.

We now establish a well-known selection principle.

THEOREM 3. *In order that $\{g_n\}$ contain a subsequence $\{g'_n\}$ such that $\{(f, g'_n)\}$ converges for each f in $\mathfrak{L}$, it is necessary and sufficient that $G = \liminf_{n\to\infty} |g_n| < +\infty$.*

The necessity of the stated condition follows at once from Theorem 1. The sufficiency is proved by a quite familiar argument which we repeat only in outline. By restricting attention to a suitable subsequence and renumbering it, we may suppose that $|g_n| \leqq K$, where K is a fixed but arbitrary constant exceeding G. The inequality $|(g_m, g_n)| \leqq |g_m|\,|g_n| \leqq K^2$ then permits us to apply the diagonal process so as to obtain a subsequence $\{g'_n\}$ for which the sequences $\{(g_m, g'_n)\}$, $m = 1, 2, 3, \cdots$ are all convergent. This is the desired sequence. Dropping primes, we have $|g_n| \leqq K$ and $\{(g_m, g_n)\}$ convergent for every m. If f is an arbitrary element in $\mathfrak{L}$, we can write $f = f_1 + f_2$, where f_1 is in the closed linear manifold $\mathfrak{M}$ generated by $\{g_n\}$ and f_2 is orthogonal to $\mathfrak{M}$. Thus, if g is an arbitrary element of $\mathfrak{L}$, we have

$$|(f, g_m) - (f, g_n)| = |(f_1, g_m) - (f_1, g_n)|$$

$$\leqq |(f_1 - g, g_m)| + |(f_1 - g, g_n)| + |(g, g_m) - (g, g_n)|$$

$$\leqq 2K|f_1 - g| + |(g, g_m) - (g, g_n)|.$$

Here we can choose g as a linear combination of $g_1, g_2, \cdots$ so that the first

[3] F. Riesz, *Zur Theorie des Hilbertschen Raumes*, Szeged Acta, vol. 7 (1934), pp. 34–38.

term is rendered small; and the second term then becomes small whenever m and n are both sufficiently great. Hence $\{(f, g_n)\}$ converges for arbitrary f in $\mathfrak{L}$.

We can restate the results of Theorems 1–3 in other terms by introducing the following definitions: a sequence $\{g_n\}$ is said to converge *weakly* or to be *weakly* convergent if $\{(f, g_n)\}$ converges for each f in $\mathfrak{L}$; to have a *weak* limit g if $(f, g_n) \to (f, g)$ for each f in $\mathfrak{L}$; to be *weakly* bounded if $\{(f, g_n)\}$ is bounded for each f in $\mathfrak{L}$; and to be bounded if $\{ | g_n | \}$ is bounded. In fact, it is easily seen that our theorems imply the following propositions:

(1) a sequence is weakly bounded (if and) only if it is bounded;
(2) $\mathfrak{L}$ is weakly complete—that is, every weakly convergent sequence has a weak limit in $\mathfrak{L}$, necessarily unique;
(3) every closed linear manifold $\mathfrak{M}$ is weakly closed and weakly complete—that is, a weakly convergent sequence in $\mathfrak{M}$ has a weak limit in $\mathfrak{M}$;
(4) the "sphere" $K(R)$ specified by the inequality $| g | \leqq R$ is weakly closed and weakly compact—that is, every sequence in $K(R)$ contains a weakly convergent subsequence with weak limit in $K(R)$.

On the other hand, Theorems 1–3 can be deduced from these four propositions in an obvious way.

We turn now to a problem in operator-theory. We prove[4]

THEOREM 4. *If the operator A has $\mathfrak{L}$ as its domain, then the following properties are equivalent:*[5]

(1) *A is a closed linear operator;*
(2) *the domain $\mathfrak{D}(A^*)$ of the adjoint A^* of A is everywhere dense in $\mathfrak{L}$;*
(3) *A is bounded and linear.*

When A has any of these three properties, $\mathfrak{D}(A^*)$ coincides with $\mathfrak{L}$ and A^* has the same bound as A.

We establish the implications (1) → (2), (2) → (3), (3) → (1), proving incidentally that (2) implies $\mathfrak{D}(A^*) = \mathfrak{L}$.

Assuming (1), we form the graph $\mathfrak{G}(A)$ of A in the direct sum[6] $\mathfrak{L} \oplus \mathfrak{L}$ and apply reasoning due to von Neumann.[7] $\mathfrak{G}(A)$ is the set of all elements of the form $\{f, Af\}$, where f belongs to $\mathfrak{L}$, the domain of A. Since A is closed and linear, $\mathfrak{G}(A)$ is a closed linear manifold. In fact, the identity

$$| f_n - f |^2 + | Af_n - f^* |^2 = | \{f_n, Af_n\} - \{f, f^*\} |^2$$

shows that $\{f_n, Af_n\} \to \{f, f^*\}$ if and only if $f_n \to f$, $Af_n \to f^*$; and the assumption that A is closed permits us to conclude the relation $f^* = Af$ or the equivalent

[4] J. v. Neumann, *Über adjungierte Funktionaloperatoren*, Annals of Mathematics, vol. 33 (1932), pp. 294–310, especially Satz 12, p. 310. See also reference to Tamarkin in footnote on p. iv of the foreword to Stone, *Linear Transformations in Hilbert Space*, New York, 1932.

[5] For definitions of terms used, see Stone, loc. cit.

[6] Stone, loc. cit., p. 30.

[7] J. v. Neumann, loc. cit.

relation $\{f, f^*\} \epsilon \mathfrak{G}(A)$. Similarly, the assumption that A is linear leads to the relations

$$\sum_{k=1}^{n} a_k\{f_k, Af_k\} = \left\{\sum_{k=1}^{n} a_k f_k, \sum_{k=1}^{n} a_k A f_k\right\} = \left\{\sum_{k=1}^{n} a_k f_k, A\left(\sum_{k=1}^{n} a_k f_k\right)\right\} \epsilon \mathfrak{G}(A).$$

The orthogonal complement of $\mathfrak{G}(A)$ is a closed linear manifold $\mathfrak{G}^*(A)$, and since $\mathfrak{G}(A)$ is a closed linear manifold, it coincides with the orthogonal complement of $\mathfrak{G}^*(A)$. Now the identity

$$(Af, g) - (f, g^*) = (\{f, Af\}, \{-g^*, g\})$$

shows that $\{-g^*, g\} \epsilon \mathfrak{G}^*(A)$ if and only if $(Af, g) = (f, g^*)$ for every f in $\mathfrak{L}$—in other words, if and only if $g \epsilon \mathfrak{D}(A^*)$ and $g^* = A^*g$. Since $\mathfrak{D}(A^*)$ is obviously a linear manifold, we can prove that $\mathfrak{D}(A^*)$ is everywhere dense in $\mathfrak{L}$ by showing that the only element h orthogonal to $\mathfrak{D}(A^*)$ is the element $h = 0$. If h is orthogonal to $\mathfrak{D}(A^*)$, then

$$(\{0, h\}, \{-g^*, g\}) = (h, g) = 0$$

for all elements $\{-g^*, g\}$ in $\mathfrak{G}^*(A)$. Hence $\{0, h\}$ is an element of $\mathfrak{G}(A)$, the orthogonal complement of $\mathfrak{G}^*(A)$, and the relation $h = A0 = 0$ is valid. We have thus deduced (2) from (1).

Now let us assume (2). If g is an arbitrary element of $\mathfrak{L}$, we choose a sequence $\{g_n\}$ in $\mathfrak{D}(A^*)$ such that $g_n \to g$. Since the relations $(f, A^*g_n) = (Af, g_n) \to (Af, g)$ hold for each f in $\mathfrak{L}$, an application of Theorem 2 establishes the existence of a unique element g^* such that $(Af, g) = (f, g^*)$ for every f in $\mathfrak{L}$. Hence $g \epsilon \mathfrak{D}(A^*)$ and $A^*g = g^*$. Our first consequence of (2) is therefore the identity $\mathfrak{D}(A^*) = \mathfrak{L}$. Since A and A^* are both defined over $\mathfrak{L}$, it is evident that the adjoint A^{**} of A^* coincides with A and that A^* shares property (2) with A. The identity $A = A^{**}$ shows that A is linear. If it were not bounded, there would exist a sequence $\{g_n\}$ with the properties

$$\liminf_{n\to\infty} |g_n| < +\infty, \qquad \lim_{n\to\infty} |Ag_n| = +\infty;$$

indeed, the first of these properties could be replaced by the stronger property $|g_n| = 1$. Now the equation $(f, Ag_n) = (A^*f, g_n)$ holds for all f in $\mathfrak{L}$. Hence we could apply the *sufficient* condition of Theorem 3 to the right-hand member, the *necessary* condition of the same theorem to the left-hand member, so as to infer the relation $\liminf_{n\to\infty} |Ag_n| < +\infty$. We would thus obtain a contradiction. Hence A must be bounded. Since our argument can be applied also to A^*, the latter operator is also bounded. If a is the bound of A—namely, the least non-negative real number such that $|Af| \leqq a|f|$ for every f in $\mathfrak{L}$—and if a^* is the bound of A^*, then the inequalities

$$|Af|^2 = (Af, Af) = (A^*Af, f) \leqq |A^*Af||f| \leqq a^*|Af||f|, |Af| \leqq a^*|f|$$

show that $a \leqq a^*$. By symmetry we have also $a^* \leqq a$ and hence $a = a^*$. Thus

we have proved not only that (3) is a consequence of (2), but also that, if A has property (2), A^* is defined over $\mathfrak{H}$ and is a bounded linear operator with the same bound as A.

Finally, the implication (3) $\rightarrow$ (1) is trivial, for if A is a bounded linear operator with $\mathfrak{H}$ as its domain and a as its bound, then the relation $f_n \rightarrow f$ implies

$$|Af_n - Af| = |A(f_n - f)| \leqq a|f_n - f| \rightarrow 0, \qquad Af_n \rightarrow Af,$$

and A is therefore closed.

A very useful application of this theorem is the following result due to Hellinger-Toeplitz.[8]

Theorem 5. *The following properties of an infinite matrix* $\{a_{mn}\}$, $m, n = 1, 2, 3, \cdots$, *of complex elements are equivalent:*

(1) $\{a_{mn}\}$ *is bounded—that is, there exists a non-negative constant* a *such that*

$$\sum_{\mu=1}^{N} \left| \sum_{\nu=1}^{N} a_{\mu\nu} x_\nu \right|^2 \leqq a^2 \sum_{\nu=1}^{N} |x_\nu|^2$$

for arbitrary complex numbers $x_1, \cdots, x_N$ *and arbitrary integral* N;

(2) *the convergence of* $\sum_{\nu=1}^{\infty} |x_\nu|^2$ *implies the convergence of*

$$\sum_{\mu=1}^{\infty} \left| \sum_{\nu=1}^{\infty} a_{\mu\nu} x_\nu \right|^2;$$

(3) *the convergence of* $\sum_{\nu=1}^{\infty} |x_\nu|^2$ *and* $\sum_{\mu=1}^{\infty} |y_\mu|^2$ *implies the convergence of the double series*

$$\sum_{\mu,\nu=1}^{\infty} a_{\mu\nu} x_\nu \bar{y}_\mu;$$

(4) *the convergence of* $\sum_{\nu=1}^{\infty} |x_\nu|^2$ *and* $\sum_{\mu=1}^{\infty} |y_\mu|^2$ *implies the convergence of the series* $\sum_{\mu=1}^{\infty} \left(\sum_{\nu=1}^{\infty} a_{\mu\nu} x_\nu \right) \bar{y}_\mu$.

These properties are equivalent to the corresponding properties for the adjoint matrix $\{a^*_{mn}\}$, *where* $a^*_{mn} = \overline{a_{nm}}$. *When any of these equivalent properties holds, we have*

$$\sum_{\mu,\nu=1}^{\infty} a_{\mu\nu} x_\nu \bar{y}_\mu = \sum_{\mu=1}^{\infty} \left(\sum_{\nu=1}^{\infty} a_{\mu\nu} x_\nu \right) \bar{y}_\mu = \sum_{\nu=1}^{\infty} x_\nu \overline{\left(\sum_{\mu=1}^{\infty} \overline{a_{\mu\nu}} y_\mu \right)}.$$

We shall establish the implications

$$\begin{array}{c} \nearrow (2) \searrow \\ (1) \rightarrow (3) \rightarrow (1) \\ \searrow (4) \rightarrow (2) \end{array}$$

[8] Hellinger-Toeplitz, Mathematische Annalen, vol. 69 (1910), pp. 289–330, esp. pp. 321–322.

and shall show further that (1) implies the final statement of the theorem. Since we have

$$\sum_{\mu,\nu=1}^{\infty} a^*_{\nu\mu} y_\nu \overline{x_\mu} = \overline{\sum_{\mu,\nu=1}^{\infty} a_{\mu\nu} x_\nu \overline{y_\mu}},$$

we see that property (3) holds for a matrix if and only if it holds for the adjoint matrix. Thus the equivalent properties (1)–(4) are equivalent to the corresponding properties for the adjoint matrix.

We shall obtain proofs by applying Theorems 1, 2 and 4 in the concrete Hilbert space $\mathfrak{H}_0$ of all sequences $x = \{x_1, x_2, x_3, \cdots\}$ for which $\sum_{\nu=1}^{\infty} |x_\nu|^2$ converges. In $\mathfrak{H}_0$, the elements $\xi_n = \{\delta_{1n}, \delta_{2n}, \delta_{3n}, \cdots\}$, $\delta_{mn} = 0$ for $m \neq n$ and $\delta_{nn} = 1$, where $n = 1, 2, 3, \cdots$, constitute a complete orthonormal set. Hence the linear manifold $\mathfrak{M}$ which they generate is everywhere dense in $\mathfrak{H}_0$. We denote by A_N the operator which carries $x = \{x_1, x_2, x_3, \cdots\}$ into $A_N x = \left\{\sum_{\nu=1}^{N} a_{1\nu}x_\nu, \cdots, \sum_{\nu=1}^{N} a_{N\nu}x_\nu, 0, 0, \cdots\right\}$. It is evident that A_N is a linear operator with $\mathfrak{H}_0$ as its domain, and it is easily shown that A_N is bounded and has an adjoint A_N^*.

Assuming (1), we first observe that, by virtue of the inequalities

$$|A_N x|^2 = \sum_{\mu=1}^{N} \left| \sum_{\nu=1}^{N} a_{\mu\nu} x_\nu \right|^2 \leqq a^2 \sum_{\nu=1}^{N} |x_\nu|^2 \leqq a^2 \sum_{\nu=1}^{\infty} |x_\nu|^2 = a^2 |x|^2$$

the operators A_N, $N = 1, 2, 3, \cdots$, are uniformly bounded. Since $(A_N \xi_n, \xi_m) = a_{mn}$ for $N \geqq m$ and $N \geqq n$, the sequence $\{(A_N x, y)\}$ converges for all x, y in $\mathfrak{M}$. Now if x, x', y, y' are arbitrary elements of $\mathfrak{H}_0$, we have

$$\begin{aligned} |(A_M x, y) - (A_N x, y)| &\leqq |(A_M(x - x'), y)| + |(A_N(x - x'), y)| \\ &\quad + |(A_M x', y - y')| + |(A_N x', y - y')| + |(A_M x', y') - (A_N x', y')| \\ &\leqq 2a|x - x'||y| + 2a|x'||y - y'| + |(A_M x', y') - (A_N x', y')|. \end{aligned}$$

We can render the first two terms in the final expression small by choosing x' and y' in $\mathfrak{M}$ so as to approximate x and y respectively; and we see that the third then becomes small for all sufficiently great M and N. The sequence $\{(A_N x, y)\}$ therefore converges for all x, y in $\mathfrak{H}_0$. Since A_N has adjoint A_N^* with $\mathfrak{H}_0$ as its domain, the equation $(A_N x, y) = (x, A_N^* y)$ shows that the sequence $\{(x, A_N^* y)\}$ is likewise convergent. According to Theorem 2, there exist unique elements Ax, A^*y, such that

$$(A_N x, y) \to (Ax, y) = (x, A^* y)$$

for all x, y in $\mathfrak{H}_0$. The operators A and A^* so defined have $\mathfrak{H}_0$ as their common domain, are both linear, and are adjoints of one another. According to Theorem 4, we see that in addition A and A^* are both bounded. We note in particular the relations

$$(A\xi_n, \xi_m) = a_{mn}, \qquad (A^*\xi_n, \xi_m) = (\xi_n, A\xi_m) = \overline{a_{nm}} = a^*_{mn}.$$

Writing $x = \sum_{\nu=1}^{\infty} (x, \xi_\nu)\xi_\nu = \sum_{\nu=1}^{\infty} x_\nu \xi_\nu$ for the arbitrary element $x = \{x_1, x_2, x_3, \cdots\}$ and applying the linearity and continuity properties of A, we now have

$$\begin{aligned} |Ax|^2 &= \sum_{\mu=1}^{\infty} |(Ax, \xi_\mu)|^2 = \sum_{\mu=1}^{\infty} \left| \left(A\left(\sum_{\nu=1}^{\infty} x_\nu \xi_\nu\right), \xi_\mu \right) \right|^2 \\ &= \sum_{\mu=1}^{\infty} \left| \sum_{\nu=1}^{\infty} (A\xi_\nu, \xi_\mu) x_\nu \right|^2 = \sum_{\mu=1}^{\infty} \left| \sum_{\nu=1}^{\infty} a_{\mu\nu} x_\nu \right|^2. \end{aligned}$$

Thus (1) implies (2). Also, putting $y = \sum_{\mu=1}^{\infty} y_\mu \xi_\mu$, we find similarly that

$$\begin{aligned} (A_N x, y) &= \left(\sum_{\mu=1}^{N} \left(\sum_{\nu=1}^{N} a_{\mu\nu} x_\nu \right) \xi_\mu, y \right) = \sum_{\mu,\nu=1}^{N} a_{\mu\nu} x_\nu (\xi_\mu, y) \\ &= \sum_{\mu,\nu=1}^{N} a_{\mu\nu} x_\nu \overline{y_\mu} \to (Ax, y). \end{aligned}$$

Thus (1) implies (3). By analogous reasoning we have

$$(Ax, y) = \sum_{\mu=1}^{\infty} (Ax, \xi_\mu)(\xi_\mu, y) = \sum_{\mu=1}^{\infty} \left(\sum_{\nu=1}^{\infty} a_{\mu\nu} x_\nu \right) \overline{y_\mu},$$

$$\begin{aligned} (Ax, y) &= (x, A^*y) = \sum_{\nu=1}^{\infty} (x, \xi_\nu)(\xi_\nu, A^*y) = \sum_{\nu=1}^{\infty} x_\nu \overline{(A^*y, \xi_\nu)} \\ &= \sum_{\nu=1}^{\infty} x_\nu \left(\sum_{\mu=1}^{\infty} \overline{a_{\mu\nu}} y_\mu \right). \end{aligned}$$

Thus (1) implies (4); and also implies the equality of the sums of the three series expressing (Ax, y). If we denote by E_N the projection of $\mathfrak{H}_0$ on the (closed) linear manifold generated by $\xi_1, \cdots, \xi_N$, we see that $A_N = E_N A E_N$. It is thus easy to show that $A_N x$ converges in $\mathfrak{H}_0$ to Ax, and that the bound of A is the least constant a for which the inequality of (1) holds.

Next let us assume (2). We can then define an operator A which carries $x = \{x_1, x_2, x_3, \cdots\}$ into $Ax = \left\{ \sum_{\nu=1}^{\infty} a_{1\nu} x_\nu, \sum_{\nu=1}^{\infty} a_{2\nu} x_\nu, \cdots \right\}$. It is evident that A is a linear operator with $\mathfrak{H}_0$ as its domain. Since

$$(x, A_N^* \xi_m) = (A_N x, \xi_m) = \sum_{\nu=1}^{N} a_{m\nu} x_\nu \to \sum_{\nu=1}^{\infty} a_{m\nu} x_\nu = (Ax, \xi_m)$$

when N becomes infinite through values exceeding m, we can apply Theorem 2 to write $(Ax, \xi_m) = (x, \xi_m^*)$ for $m = 1, 2, 3, \cdots$. Hence A has adjoint A^* defined throughout the linear manifold $\mathfrak{M}$. By Theorem 4, the operator A is bounded. If a is the bound of A, we have

$$\sum_{\mu=1}^{\infty}\left|\sum_{\nu=1}^{\infty} a_{\mu\nu}x_\nu\right|^2 = |Ax|^2 \leqq a^2|x|^2 = a^2\sum_{\nu=1}^{\infty}|x_\nu|^2.$$

Applying this inequality to the case where $x = \{x_1, x_2, \cdots, x_N, 0, 0, \cdots\}$, we have

$$\sum_{\mu=1}^{N}\left|\sum_{\nu=1}^{N} a_{\mu\nu}x_\nu\right|^2 \leqq \sum_{\mu=1}^{\infty}\left|\sum_{\nu=1}^{N} a_{\mu\nu}x_\nu\right|^2 \leqq a^2\sum_{\nu=1}^{N}|x_\nu|^2.$$

Hence (2) implies (1). We may remark that the operator A defined in the present paragraph is now easily identified with the operator A introduced in the preceding one, by virtue of the relation $(A_N x, \xi_m) \to (Ax, \xi_m)$ which holds for *both* operators A.

Next let us assume (3). If $x = \{x_1, x_2, x_3, \cdots\}$ and $y = \{y_1, y_2, y_3, \cdots\}$ are arbitrary elements of $\mathfrak{H}_0$, we see that

$$(A_N x, y) = \sum_{\mu,\nu=1}^{N} a_{\mu\nu}x_\nu\overline{y_\mu} = (x, A_N^* y)$$

converges. By Theorem 2 there exist unique elements Ax, A^*y such that

$$(Ax, y) = \sum_{\mu,\nu=1}^{\infty} a_{\mu\nu}x_\nu\overline{y_\mu} = (x, A^*y),$$

for all x, y in $\mathfrak{H}_0$. The operators A and A^* thus defined have $\mathfrak{H}_0$ as their common domain, are both linear, and are adjoints of one another. According to Theorem 4, they are both bounded, with common bound a. Consequently we have

$$\left|\sum_{\mu,\nu=1}^{\infty} a_{\mu\nu}x_\nu\overline{y_\mu}\right|^2 \leqq a^2|x|^2|y|^2 = a^2\sum_{\nu=1}^{\infty}|x_\nu|^2\sum_{\mu=1}^{\infty}|y_\mu|^2.$$

Applying this inequality to the case where $x = \{x_1, \cdots, x_N, 0, 0, \cdots\}$ and $y = A_N x$, we have

$$\left|\sum_{\mu=1}^{N}\left|\sum_{\nu=1}^{N} a_{\mu\nu}x_\nu\right|^2\right|^2 \leqq a^2\sum_{\nu=1}^{N}|x_\nu|^2\sum_{\mu=1}^{N}\left|\sum_{\nu=1}^{N} a_{\mu\nu}x_\nu\right|^2,$$

and hence

$$\sum_{\mu=1}^{N}\left|\sum_{\nu=1}^{N} a_{\mu\nu}x_\nu\right|^2 \leqq a^2\sum_{\nu=1}^{N}|x_\nu|^2.$$

Thus (3) implies (1). Again the operator A introduced here is easily identified with those previously defined.

Finally, we assume (4). Then for arbitrary $x = \{x_1, x_2, x_3, \cdots\}$ in $\mathfrak{H}_0$ the series $\sum_{\nu=1}^{\infty} a_{m\nu}x_\nu$ converges to a sum z_m; and for arbitrary $y = \{y_1, y_2, y_3, \cdots\}$ in $\mathfrak{H}_0$, the series $\sum_{\mu=1}^{\infty} z_\mu\overline{y_\mu}$ converges. Let ζ_N be the element $\{z_1, z_2, \cdots, z_N,$

0, 0, $\cdots\}$ in $\mathfrak{H}_0$. Then $(\zeta_N, y) = \sum_{\mu=1}^{N} z_\mu \overline{y_\mu}$ converges for arbitrary y in $\mathfrak{H}_0$. Hence Theorem 1 shows that

$$|\zeta_N|^2 = \sum_{\mu=1}^{N} |z_\mu|^2 = \sum_{\mu=1}^{N} \left| \sum_{\nu=1}^{\infty} a_{\mu\nu} x_\nu \right|^2$$

remains bounded. We see therefore that (4) implies (2).

With this the proof of our theorem is complete.

HARVARD UNIVERSITY AND BROWN UNIVERSITY.

THE STRUCTURE OF CERTAIN RATIONAL INFINITE ALGEBRAS

By Otto F. G. Schilling

G. Köthe has investigated infinite algebras over abstract fields in his paper "Ueber Schiefkörper unendlichen Ranges".[1] In this note we shall apply some of his fundamental results to infinite algebras over a finite algebraic number field. Application of the theory of finite algebras over algebraic number fields enables us to give explicit representations of the algebras considered as generalized crossed products.[2] Furthermore, we shall investigate the arithmetic in such infinite algebras.

1. **Algebraic theory.** Let k be an algebraic number field of finite degree over the field of all rational numbers. We consider algebras of infinite rank over k as centrum. We assume that the algebras A are countable, i.e., that there exists a countable set $\{a_1, a_2, \cdots, a_i, \cdots\}$ of elements of A, such that each element a of A can be represented as a finite sum $\sum_{\nu=1}^{\nu_0} k_{i_\nu} a_{i_\nu}$ with coefficients k_{i_ν} in the field k. Furthermore, we restrict ourselves to completely normal algebras which we define as follows.

Definition. A countable infinite algebra A with centrum k is called totally normal over k if every finite system $\{b_1, \cdots, b_m\}$ of elements of A lies in a normal simple finite algebra over k.

With slight modifications of Köthe's proofs one proves

Theorem 1. *For every totally normal algebra A there exists at least one defining sequence* $k \subseteq \cdots \subseteq A_{i-1} \subseteq A_i \subseteq \cdots$ *of normal simple algebras* A_i.

Conversely, we have

Theorem 2. *Every sequence* $k \subseteq \cdots \subseteq A_{i-1} \subseteq A_i \subseteq \cdots$ *of normal simple algebras* A_i *over k defines an infinite totally normal algebra A with the center k.*

Theorem 3. *Every totally normal algebra A over k is representable as the direct product of a countable infinity of normal simple algebras* A_i *of finite degrees over k. Such a decomposition is not necessarily uniquely determined.*

If we collect all simple normal systems of such a decomposition which belong to a fixed prime q, we have

Theorem 4. *Every totally normal algebra A over k is the direct product of a countable infinity of simple algebras* $A^{(q)}$ *which are primary with respect to k. The factors are uniquely determined except for isomorphisms.*

Received December 23, 1936.

[1] G. Köthe, Math. Annalen, vol. 105 (1931), pp. 15-39. See this paper also for the different notions used later on.

[2] For the theory of crossed products see the report of Max Deuring, *Algebren*, Berlin, 1935.

Thus the determination of structure is reduced for this type of algebra to that of the structure of finite or infinite totally normal primary algebras $A^{(q)}$.

Now let us consider a direct product $A = A_1 \times \cdots \times A_i \times \cdots$ of infinitely many finite primary algebras A_i. Then $A_n^* = A_1 \times \cdots \times A_n = D_n \times k_{r_n}$ with a normal division algebra D_n and total matric algebra k_{k_n} of degree r_n over the field k. Köthe has shown

THEOREM 5. *If the degrees r_n of the matric algebras successively contained in A_n^* tend to infinity with increasing n, the infinite direct product A is isomorphic to the direct product M of infinitely many matric algebras k_q.*

With the help of Theorem 3 and the theory of finite normal simple algebras over an algebraic number field k, we prove

THEOREM 6. *There do not actually exist infinite totally normal primary division algebras D over a finite number field k.*

Proof. It is a well-known fact that the direct product of two normal finite division algebras D_1 and D_2 over k is a division algebra if and only if the exponent of $D_1 \times D_2$ is equal to the degree of $D_1 \times D_2$. The degree of $D_1 \times D_2$ is equal to the product of the degrees of the factors. Now let D_1 and D_2 be primary division algebras with the respective degrees q^{α_1} and q^{α_2}, α_1, α_2 both > 0. Then the algebra $D_1 \times D_2$ has the degree $q^{\alpha_1+\alpha_2}$.

On the other hand, the index of an algebra is determined as the least common multiple of the local exponents or indices.[3] According to the theory of finite normal algebras over an algebraic number field, the local invariants of a product are equal to the sum of the respective local invariants of the factors. The latter can be represented for every prime ideal $\mathfrak{p}$ of k as fractions with the maximal denominators $q^{\max(\alpha_1, \alpha_2)}$. The possible exponent in the large is then at most $q^{\max(\alpha_1, \alpha_2)}$.[4] But this means that the direct product of two primary finite division algebras D_1 and D_2 must necessarily split off a matric algebra, because $\max(\alpha_1, \alpha_2) < \alpha_1 + \alpha_2$. We apply this to the products $D_1 \times D_2 \times \cdots \times D_n = D'_n \times k_{r_n}$, where D'_n is the division algebra which belongs to the product. Then one sees immediately that the degrees r_n of the matric algebras tend to infinity with increasing n. According to Theorem 5 all these products are isomorphic to the infinite product of matric algebras k_q. Every division algebra D can be represented as an infinite direct product of finite division algebras as stated in Theorem 3. But our investigations of direct products of primary division algebras show that they must be isomorphic to M.

There remains then the study of totally normal algebras which are not isomorphic to an M. They must be representable as infinite direct products of primary finite normal simple algebras, as Theorems 3 and 6 show.

We apply the theory of structure as developed by H. Hasse[5] to this problem. He shows that every finite normal algebra A over the algebraic number field k

[3] Cf. footnote 2.

[4] Cf. footnote 2.

[5] H. Hasse, *Die Struktur der R. Brauerschen Algebrenklassengruppe über einem algebraischen Zahlkörper*, Math. Annalen, vol. 107 (1933), pp. 731–760.

of degree n can be represented as a cyclic crossed product $(a, Z/k, u)$, where Z is cyclic over k of degree n and a is an element $(\neq 0)$ of k, u an operator, describing the algebra A with respect to the splitting field Z. The complete factor set of this representation is given by the square matrix of n rows and columns

$$F^{(n)}(a) = \begin{bmatrix} 1 & \cdots & \cdots & 1 \\ \vdots & & \diagup & a \\ \vdots & \diagup & \diagup & \vdots \\ 1 & a & \cdots & a \end{bmatrix}.$$

Therefore we may write $A = (F^{(n)}(a), Z/k, u)$.

Now we consider two prime[6] cyclic products $A_1 = (F^{(n_1)}(a_1), Z_1/k, u_1)$ and $A_2 = (F^{(n_2)}(a_2), Z_2/k, u_2)$. Their direct product is isomorphic to a cyclic algebra with the splitting field Z_1Z_2. This field Z_1Z_2 is cyclic over k, because we assumed A_1 prime to A_2. The Galois group of Z_1Z_2/k is isomorphic to the direct product of the Galois groups $\{S_1\}$ and $\{S_2\}$ of Z_1/k and Z_2/k. For this reason we shall represent $A_1 \times A_2$ by the symbol

$$(F^{(n_1)}(a_1) \otimes F^{(n_2)}(a_2), \qquad Z_1Z_2/k, \qquad \{u_1, u_2\}),$$

where $F^{(n_1)}(a_1) \otimes F^{(n_2)}(a_2)$ is an abbreviation for the factor set arising from the $(n_1n_2)^2$ products

$$u_1^{m_1} \cdot u_2^{m_2} = a_1^{g_1} a_2^{g_2} \cdot u_1^{\mu_1} \cdot u_2^{\mu_2},$$

if $m_1 = g_1n_1 + \mu_1$ and $m_2 = g_2n_2 + \mu_2$. This factor set $F^{(n_1)}(a_1) \otimes F^{(n_2)}(a_2)$ has the form

$$\begin{pmatrix} F^{(n_1)}(a_1) & \cdots & & F^{(n_1)}(a_1) \\ \vdots & & & a_2 \cdot F^{(n_1)}(a_1) \\ \vdots & & & \vdots \\ F^{(n_1)}(a_1) & a_2 \cdot F^{(n_1)}(a_1) & \cdots & a_2 \cdot F^{(n_1)}(a_1) \end{pmatrix},$$

with n_2 rows and columns. It is the multiplication table belonging to the n_1n_2 elements of the complex $\{u_1, u_2\}$ in the ordering

$$1, u_1, u_1^2, \cdots, u_1^{n_1-1}, \qquad u_2, u_2^2, \cdots, u_2^{n_2-1}, \qquad u_1u_2, u_1^2u_2, \cdots, u_1u_2^2, \qquad \cdots.$$

The matrix $F^{(n_2)}(a_2) \otimes F^{(n_1)}(a_1)$ is generated as the tableau of

$$1, u_2, u_2^2, \cdots, u_2^{n_2-1}, \qquad u_1, u_1^2, \cdots, u_1^{n_1-1}, \qquad u_2u_1, u_2^2u_1, \cdots, u_2u_1^2, \qquad \cdots.$$

Obviously $F^{(n_2)}(a_2) \otimes F^{(n_1)}(a_1)$ can be obtained from $F^{(n_1)}(a_1) \otimes F^{(n_2)}(a_2)$ by interchange of rows and columns, and vice versa.

Furthermore, it is required that $u_2^{-1}z_1u_2 = z_1$ and $u_1^{-1}z_2u_1 = z_2$ for arbitrary elements z_1, z_2 $(\neq 0)$ in Z_1 and Z_2. We shall not reduce this representation of $A_1 \times A_2$ to the cyclic form $(F^{(n_1 n_2)}(a_{12}), Z_1Z_2/k, u_{12})$.

If we form the union Z of infinitely many finite algebraic number fields Z_i

[6] Two cyclic products are said to be prime if their degrees are prime.

which are prime over k and cyclic, this field Z is an infinite cyclic extension of k. The direct product $\mathbf{z}$ of all the cyclic Galois groups of the fields Z_i/k is a group of automorphisms of the field Z/k. It is contained in the most general group of automorphisms of Z/k as introduced by J. Herbrand[7] and W. Krull[8] by topological methods. All elements of $\mathbf{z}$ are of finite order by definition of the direct product.

Now let us consider infinitely many crossed products $A_i = (F^{(n_i)}(a_i), Z_i/k, u_i)$, which are prime with respect to k. Then the infinite complex $\{u_1, u_2, \cdots\}$ with the relations

$$u_{i_1}^{m_{i_1}} \cdot \cdots \cdot u_{i_j}^{m_{i_j}} = u_{i_1}^{\mu_{i_1}} \cdot \cdots \cdot u_{i_j}^{\mu_{i_j}} a_{i_1}^{g_{i_1}} \cdot \cdots \cdot a_{i_j}^{g_{i_j}}$$

with $m_{i_j} = g_{i_j} n_{i_j} + \mu_{i_j}$ for every finite product of operators forms a crossed representation of the group $\mathbf{z}$. The factor set arising above may be called $F^{(n_1)}(a_1) \otimes F^{(n_2)}(a_2) \otimes \cdots$. We define then the infinite algebra

$$A = (F^{(n_1)}(a_1) \otimes F^{(n_2)}(a_2) \otimes \cdots, Z/k, \{u_1, u_2, \cdots\})$$

as the set of all finite sums $\sum_{\sigma, \tau} k_{\sigma_1 \cdots \sigma_s \tau_1 \cdots \tau_t} z_{\sigma_1} \cdots z_{\sigma_s} u_{\tau_1}^{\rho_{\tau_1}} \cdots u_{\tau_t}^{\rho_{\tau_t}}$ with elements $k\ldots$ in the field k.

The laws of composition are completely defined by the relations

$$\text{(i)} \qquad \prod_{\nu=1}^{j} u_{i_\nu}^{m_{i_\nu}} = \prod_{\nu=1}^{j} a_{i_\nu}^{g_{i_\nu}} \cdot \prod_{\nu=1}^{j} u_{i_\nu}^{\mu_{i_\nu}}$$

$$\text{(ii)} \qquad \begin{aligned} u_i^{-1} z_j u_i &= z_j \qquad (i \neq j) \text{ for every } z_j \neq 0 \text{ of } Z_j; \\ u_i^{-1} z_i u_i &= z_i^{S_i} \text{ for every } z_i \neq 0 \text{ of } Z_i. \end{aligned}$$

We shall call such an infinite crossed product with cyclic Z/k a generalized cyclic product. Then we have

THEOREM 7. *The generalized cyclic products A are totally normal over the field k.*

Proof. It is obvious that the finite direct products $A_1 \times A_2 \times \cdots \times A_i$ form a defining totally normal sequence of A. Theorem 2 asserts that A is then totally normal.

THEOREM 8. *Every proper[9] infinite totally normal algebra A over k can be represented as a generalized cyclic product.*

Proof. Let A be decomposed according to Theorem 4. The components are finite simple normal algebras A_i which are prime with respect to k. Theorem 7 leads to possible representations. We remark that these representations are not at all uniquely determined.

[7] J. Herbrand, *Extensions algébriques de degré infini*, Math. Annalen, vol. 108 (1933), pp. 699–717.

[8] W. Krull, *Galoissche Theorie der unendlichen algebraischen Erweiterungen*, Math. Annalen, vol. 100 (1928), pp. 687–698.

[9] That is to say an algebra which is not isomorphic to an infinite matric algebra M.

2. **Arithmetic theory.** We first introduce some notions already known in the theory of infinite algebraic number fields.[10] Let $k = A_0 \subseteq \cdots \subseteq A_{i-1} \subseteq A_i \subseteq \cdots$ be a defining sequence of normal simple systems of the totally normal infinite algebra A over k.

DEFINITION. We say that the series $\mathbf{R}_0 \subseteq \cdots \subseteq \mathbf{R}_{i-1} \subseteq \mathbf{R}_i \subseteq \cdots$ of rings $\mathbf{R}_i$ in A_i is a sequence "belonging to $A = \{A_i\}$" if and only if

(i) the rings $\mathbf{R}_i$ are of maximal rank in A_i,

(ii) $A_{i-1} \cap \mathbf{R}_i = \mathbf{R}_{i-1}$.

We consider also ideals $\mathbf{a}_i$ in the rings $\mathbf{R}_i$: left, right, and two-sided ideals.

DEFINITION. A series $\mathbf{a}_0 \subseteq \cdots \subseteq \mathbf{a}_{i-1} \subseteq \mathbf{a}_i \subseteq \cdots$ of ideals is called a sequence of ideals belonging to the sequence of rings $\mathbf{R}_i$ if and only if

(i) $\mathbf{a}_i \subseteq \mathbf{R}_i, \quad \mathbf{a}_i\mathbf{R}_i \subseteq \mathbf{a}_i$,

(ii) $\mathbf{R}_{i-1} \cap \mathbf{a}_i = \mathbf{a}_{i-1}$.

Condition (i) is to be changed according to the nature of the ideals.

Then we have

THEOREM 8. *Between the rings* $\mathbf{R}$ *of* A *and the sequences of rings* $\{\mathbf{R}_i\}$ *can be established a one-to-one correspondence, which is explicitly given by*

$$\{\mathbf{R}_i\} \to \mathbf{R} = \Sigma \mathbf{R}_i, \qquad \mathbf{R} \to \{\mathbf{R} \cap A_i = \mathbf{R}_i\}.$$

THEOREM 9. *Between the ideals* $\mathbf{a}$ *of the rings* $\mathbf{R}$ *in* A *and the sequences of ideals* $\{\mathbf{a}_i\}$ *belonging to the sequence of rings* $\mathbf{R} = \{\mathbf{R}_i\}$ *there holds a one-to-one correspondence according to the formulas*

$$\{\mathbf{a}_i\} \to \mathbf{a} = \Sigma \mathbf{a}_i, \qquad \mathbf{a} \to \{\mathbf{a} \cap \mathbf{R}_i = \mathbf{a}_i\}.$$

For the proofs one has only to recall the different definitions.

DEFINITION. A ring $\mathbf{J}$ of A is called a maximal order if

(i) $\mathbf{J}$ contains the maximal order of k,

(ii) all elements of $\mathbf{J}$ satisfy minimal equations with coefficients in the maximal order of k,

(iii) $\mathbf{J}$ is not contained in a larger ring, which fulfills the conditions (i) and (ii).

We now construct special maximal orders of A. Let $A = A_1 \times A_2 \times \cdots \times A_i \times \cdots$ be a representation of A according to Theorem 3. In each system A_i we fix an arbitrary maximal order $\mathbf{M}_i$. Then we form successively the direct products

$$\mathbf{M}_1 \times \mathbf{M}_2 \subseteq \mathbf{M}_{12} \subset A_1 \times A_2$$

$$\mathbf{M}_{12} \times \mathbf{M}_3 \subseteq \mathbf{M}_{123} \subset A_1 \times A_2 \times A_3$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\mathbf{M}_{12\cdots i-1} \times \mathbf{M}_i \subseteq \mathbf{M}_{12\cdots i} \subset A_1 \times A_2 \times \cdots \times A_i.$$

[10] W. Krull, *Idealtheorie in unendlichen Zahlkörpern*, Math. Zeitschrift, vol. 29 (1929), pp. 12–54.

We denote by $\mathbf{M}_{12\cdots i}$ the *uniquely* determined[11] maximal order of $A_1 \times A_2 \times \cdots \times A_i$ which contains the order $\mathbf{M}_{12\cdots i-1} \times \mathbf{M}_i$. All these hypercomplex orders have the highest possible rank.

We proceed to

THEOREM 10. *The sequence of maximal orders* $\{\mathbf{M}_{12\cdots i}\}$ *defines a maximal order* $\mathbf{J}$ *of* A.

Proof. Let us assume that $\mathbf{J}$ is not a maximal order. Then $\mathbf{J}$ is contained in at least one maximal order $\mathbf{J}'$. Hence there must exist an element a of $\mathbf{J}'$ which does not lie in $\mathbf{J}$. Since a is an element of a maximal order $\mathbf{J}'$, its minimal equation has integer coefficients in k. It must belong to a finite subalgebra of A. Let $A_1 \times A_2 \times \cdots \times A_n$ be this algebra. Then $\mathbf{M}' = \{\mathbf{J}, a\} \cap A_1 \times A_2 \times \cdots \times A_n \supset \mathbf{M}_{1\cdots n}$. This means that in the finite normal algebra $A_1 \times A_2 \times \cdots \times A_n$ there must exist an order $\mathbf{M}'$ properly containing the maximal order $\mathbf{M}_{1\cdots n}$, but this is clearly impossible.

THEOREM 11. *There exist continuously many maximal orders in the algebra* A.

Proof. This follows immediately by considering the construction of the special maximal order $\mathbf{J}$. At each step we can choose infinitely many maximal orders $\mathbf{M}_i$ in the algebras A_i to form the maximal orders $\mathbf{M}_{12\cdots i}$ of an approximation. The maximal orders $\mathbf{M}_{12\cdots i}$ are different for different $\mathbf{M}_{i-1}$. To prove this let $\mathbf{M}_i$ and $\mathbf{M}'_i$ be two different maximal orders of A_i. Assume that the uniquely determined maximal orders $\mathbf{M}_{12\cdots i}$ and $\mathbf{M}'_{12\cdots i}$ in $A_1 \times \cdots \times A_i$ which imbed $\mathbf{M}_{i-1}$ are equal. Then the intersections $\mathbf{M}_{12\cdots i} \cap A_1 \times \cdots \times A_{i-1}$ and $\mathbf{M}'_{12\cdots i} \cap A_1 \times \cdots \times A_{i-1}$ are equal. This intersection is obviously an order of $A_1 \times \cdots \times A_{i-1}$. It contains the two different maximal orders $\mathbf{M}_i$ and $\mathbf{M}'_i$. It therefore is equal to both, in contradiction to the assumption. This construction leads then to $\aleph_0^{\aleph_0} = \aleph$ different maximal orders of A.

Now we consider the decomposition of prime ideals $\mathfrak{p}$ of the center k. It suffices to consider the $\mathfrak{p}$-adic extensions $A_\mathfrak{p}$ of the given totally normal algebra A.[12] We understand that the extension $A_\mathfrak{p}$ is the product modulus $A \cdot k_\mathfrak{p}$. Obviously $A_\mathfrak{p}$ is a totally normal simple algebra with the center $k_\mathfrak{p}$.[13] In the representation $(A_1)_\mathfrak{p} \times \cdots \times (A_i)_\mathfrak{p} \times \cdots$ of $A_\mathfrak{p}$ as a direct product, we can combine all finite local division algebras $(D_\mathfrak{p})_i$ of $(A_i)_\mathfrak{p}$ into a local division algebra $D_\mathfrak{p}$ which is also totally normal over $k_\mathfrak{p}$.

This reduces the problem of decomposition to the two partial problems: (1) arithmetic in a totally normal division algebra $D_\mathfrak{p}$ over $k_\mathfrak{p}$, (2) arithmetic in systems of matrices over such a division algebra. We prove

THEOREM 12. *A totally normal division algebra* $D_\mathfrak{p}$ *over a* $\mathfrak{p}$*-adic number field* $k_\mathfrak{p}$ *contains exactly one maximal order* $\mathbf{O}_\mathfrak{p}$, *and there exists one prime ideal* $\mathbf{P}$ *in* $D_\mathfrak{p}$.

[11] T. Nakayama, *Über das Produkt zweier einfachen Algebren mit zu einander teilerfremden p-Indizen*, Jap. Journal of Math., vol. 12 (1935).

[12] E. Noether, *Zerfallende verschränkte Produkte und ihre Maximalordnungen*, Actualités Scientifiques et Industrielles, 1934.

[13] Consider the approximation by products of the $(A_i)_\mathfrak{p}$.

All other ideals are two-sided and they are determined by their values with respect to $\mathbf{P}$ *and their symbols.*

Proof. Let $\cdots \subseteq (D_\mathfrak{p})_{i-1} \subseteq (D_\mathfrak{p})_i \subseteq \cdots$ be a defining sequence of finite $\mathfrak{p}$-adic division algebras of degrees n_i over $k_\mathfrak{p}$. These algebras contain exactly one maximal order $(\mathbf{O}_\mathfrak{p})_i$ with one prime ideal $\mathbf{P}_i$. The value with respect to $\mathbf{P}_i$ may be denoted by $v_{\mathbf{P}_i}$.[14] The sum $\sum (\mathbf{O}_\mathfrak{p})_i$ is obviously a maximal order $\mathbf{O}_\mathfrak{p}$ of the algebra $D_\mathfrak{p}$. By definition, $\mathbf{O}_\mathfrak{p}$ contains all elements of $D_\mathfrak{p}$ whose minimal equation has integer coefficients in the field $k_\mathfrak{p}$ and the highest coefficient is one. Let $\mathbf{O}'_\mathfrak{p}$ be another maximal order of $D_\mathfrak{p}$. All its elements satisfy minimal equations with integer coefficients in $k_\mathfrak{p}$. Therefore $\mathbf{O}'_\mathfrak{p} \subseteq \mathbf{O}_\mathfrak{p}$. But $\mathbf{O}'_\mathfrak{p}$ was assumed to be a maximal order. Hence $\mathbf{O}'_\mathfrak{p} = \mathbf{O}_\mathfrak{p}$.

The sequence $\cdots \subseteq \mathbf{P}_{i-1} \subseteq \mathbf{P}_i \subseteq \cdots$ of prime ideals defines the prime ideal $\mathbf{P}$ of the maximal order $\mathbf{O}_\mathfrak{p}$. Let $\mathbf{a}$ be an ideal in $\mathbf{O}_\mathfrak{p}$. Its defining sequence $\{\mathbf{a}_i = \mathbf{a} \cap (\mathbf{O}_\mathfrak{p})_i\}$ of ideals in $(\mathbf{O}_\mathfrak{p})_i$ consists of two-sided ideals. Hence $\mathbf{a}$ itself is two-sided. The value $v_\mathbf{P}$ with respect to the prime ideal $\mathbf{P}$ is defined by $v_\mathbf{P}(a) = \lim_{i \to \infty} v_{\mathbf{P}_i}(a) \cdot n_i^{-1}$, if $\mathbf{a}$ lies in the algebra $(D_\mathfrak{p})_{i_0}$ and not in $(D_\mathfrak{p})_{i_0-1}$. It is $v_\mathfrak{p}(a) = v_{\mathfrak{p}_0}(a) \cdot n_{i_0}^{-1}$. The value $v_\mathbf{P}(\mathbf{a})$ of an ideal $\mathbf{a}$ is defined by $\liminf_{a \in \mathbf{a}} v_\mathbf{P}(a)$. The symbol $s(\mathbf{a})$ is equal to f (finite) if and only if there exists an element a in $\mathbf{a}$ such that $v_\mathbf{P}(a) = v_\mathbf{P}(\mathbf{a})$. The symbol $s(\mathbf{a}) = i$ (infinite) in all other cases. Then all ideals of $\mathbf{O}_\mathfrak{p}$ are uniquely determined by their values and symbols.[15] This theory runs exactly as in the commutative case.

If we consider a finite matric algebra $(D_\mathfrak{p})_m$ over $D_\mathfrak{p}$, then $(\mathbf{O}_\mathfrak{p})_m$ is a maximal order of $(D_\mathfrak{p})_m$.

THEOREM 13. *The two-sided prime ideal* $\overline{\mathbf{P}}$ *of* $(\mathbf{O}_\mathfrak{p})_m$ *is given by*

$$\mathbf{P}(\mathbf{O}_\mathfrak{p})_m = \begin{pmatrix} \mathbf{P} & \cdots & \mathbf{P} \\ \cdot & & \cdot \\ \mathbf{P} & \cdots & \mathbf{P} \end{pmatrix}^{(m)}$$

and it is the common part of m left ideals $\mathbf{L}_j$ *which are given in the form*

$$\mathbf{L}_j = \begin{pmatrix} \mathbf{O}_\mathfrak{p} & \cdots & \mathbf{P} & \cdots & \mathbf{O}_\mathfrak{p} \\ \cdot & & \cdot & & \cdot \\ \mathbf{O}_\mathfrak{p} & \cdots & \mathbf{P} & \cdots & \mathbf{O}_\mathfrak{p} \end{pmatrix}^{(m)},$$

the $\mathbf{P}$*'s being in column j.*

Proof. These facts are quite obvious.[16]

Now we consider the product $A_\mathfrak{p}$ of infinitely many matric algebras $(k_\mathfrak{p})_{m_i}$ with the totally normal division algebra $D_\mathfrak{p}$.

[14] For the notion of value, etc., in division algebras see H. Hasse, *Über p-adische Schiefkörper und ihre Bedeutung für die Arithmetik hyperkomplexer Zahlen*, Math. Annalen, vol. 104 (1930), pp. 495–534. See this paper also for the proofs of the following theorems.

[15] Cf. footnote 10. The algebra $D_\mathfrak{p}$ is by no means perfect with respect to the valuation $v_\mathbf{P}$.

[16] Cf. footnote 14.

THEOREM 14. *The two-sided prime ideal* $\bar{\mathbf{P}}$ *of the maximal order* $\mathbf{O}_{\mathfrak{p}} \times \Pi(\mathbf{o}_{\mathfrak{p}})_{m_i} = \mathbf{J}$ *is divisible by continuously many left ideals in* $\mathbf{J}$. ($\mathbf{o}_{\mathfrak{p}}$ *is the maximal order of* $k_{\mathfrak{p}}$.)

Proof. The series $\mathbf{o}_{\mathfrak{p}} \subseteq \cdots \subseteq \mathbf{O}_{\mathfrak{p}} \times \prod_{j=1}^{i-1} (\mathbf{o}_{\mathfrak{p}})_{m_j} \subseteq \mathbf{O}_{\mathfrak{p}} \times \prod_{j=1}^{i} (\mathbf{o}_{\mathfrak{p}})_{m_j} \subseteq \cdots$ is a defining sequence for a maximal order $\mathbf{J}$ of $A_{\mathfrak{p}}$. The ideals $\bar{\mathbf{P}} \times \mathbf{O}_{\mathfrak{p}} \times \prod_{j=1}^{i-1} (\mathbf{o}_{\mathfrak{p}})_{m_j}$ are two-sided prime ideals of the maximal orders $\mathbf{O}_{\mathfrak{p}} \times \prod_{j=1}^{i-1} (\mathbf{o}_{\mathfrak{p}})_{m_j}$. They define the two-sided prime ideal $\bar{\bar{\mathbf{P}}}$ of $\mathbf{J}$. Let $\mathbf{L}_n$ be one of the left divisors of $\mathbf{P} \times \mathbf{O}_{\mathfrak{p}} \times \prod_{j=1}^{i-1} (\mathbf{o}_{\mathfrak{p}})_{m_j}$ in $\mathbf{O}_{\mathfrak{p}} \times \prod_{j=1}^{i-1} (\mathbf{o}_{\mathfrak{p}})_{m_j}$. Their number is surely greater than one. Then all of these ideals $\mathbf{L}_{i-1}$ are divisible by more than one left ideal $\mathbf{L}_i'$ of $\mathbf{O}_{\mathfrak{p}} \times \prod_{j=1}^{i} (\mathbf{o}_{\mathfrak{p}})_{m_j}$. This construction leads again to at least $2^{\aleph_0} = \aleph$ left divisors of $\bar{\bar{\mathbf{P}}}$.

Remark. A combination of these results with the explicit construction of finite division algebras shows that there exist totally normal division algebras in which exactly two prime ideals $\mathfrak{p}$ and $\mathfrak{q}$ are totally ramified, i.e., $D_{\mathfrak{p}}$ and $D_{\mathfrak{q}}$ are both infinite division algebras. There also exist totally normal division algebras, all $\mathfrak{p}$-adic extensions of which are equivalent to the product of a finite division algebra with an infinite product of matric algebras.

THE INSTITUTE FOR ADVANCED STUDY.

AN EXTENDED ARITHMETIC

By Garrett Birkhoff

1. **Introduction.** In this paper there are defined three combinatorial operations upon partially ordered sets X and Y resulting in partially ordered sets which will be denoted by $X + Y$, XY, and X^Y respectively.

In the case where X and Y are finite unordered sets with cardinal numbers m and n respectively, the operations yield the finite unordered sets of cardinal numbers $m + n$, mn, and m^n in the commonplace sense. In the more general case where the requirement of finiteness is dropped, they yield the usual foundations for the arithmetic of general cardinal numbers.[1]

It will be proved that the formal properties of general cardinal arithmetic[2] persist as laws of composition for our extended arithmetic of general partially ordered sets. On the other hand, no algorithm for well-ordering the class of partially ordered systems is given, and so the theorem of transfinite arithmetic which asserts that the cardinal numbers are well-ordered has no analogue.

It will also be shown that our extended arithmetic is of considerable use in describing algebraic systems. Thus although it differs from Hausdorff's ordinal arithmetic when applied to sequences, it is apparently more consequential than the latter.[3]

2. **The extended arithmetic.** The extended arithmetic which is proposed will be described by defining first the domain of *elements* to which its operations apply, and then defining the resultants of its operations.

The elements are *partially ordered systems* in the usual sense of Hausdorff—that is, systems X, Y, Z, $\cdots$ whose members (denoted by small Latin letters) are related by an inclusion relation $x \leqq x'$ satisfying

P1: $x \leqq x$. (Reflexiveness)
P2: $x \leqq x'$ and $x' \leqq x$ imply $x = x'$. (Anti-symmetry)
P3: $x \leqq x'$ and $x' \leqq x''$ imply $x \leqq x''$. (Transitivity)

Two partially ordered systems X and Y will be called *isomorphic* (written

Received May 22, 1936; in revised form, December 31, 1936.

[1] Cf. F. Hausdorff's *Mengenlehre*, Berlin, 1927, p. 29, p. 62.

[2] The possibility of such relations as $x + 1 = x$, $2x = x$, and $x^2 = x$ distinguishes general from finite cardinal arithmetic.

[3] Although transfinite numbers have numerous uses, the only constructions of sums, products, or powers of transfinite numbers which have been really interesting hitherto have been those of (1) the power of the continuum as $2^{\aleph_0}$ ($\aleph_0$ denotes countable infinity), and (2) transfinite ordinals such as $\omega + 1$, 3ω, or ω^2 by ordinal addition and multiplication, and only the second of these is lost in our extended arithmetic.

$X = Y$ in this article) if and only if there exists a one-one correspondence between their elements which preserves inclusion.

By the *sum* $X + Y$ of two such systems X and Y is meant the system whose members include both the members x of X and the members y of Y, in which $x \leqq x'$ and $y \leqq y'$ keep their meaning, while $x \leqq y$ and $y \leqq x$ are always denied.

By the *product* $X \cdot Y$ of X and Y is meant the system whose members are all couples $[x, y]$ with $x \in X$ and $y \in Y$, and in which $[x, y] \leqq [x', y']$ means that $x \leqq x'$ in X and $y \leqq y'$ in Y.

By the *power* X^Y of one such system X with respect to a second such system Y as exponent is meant the system whose members are the *monotonic functions* $f(y)$ with domain Y and range in X (that is, all functions such that $y \leqq y'$ in Y implies $f(y) \leqq f(y')$ in X), ordered by having $f \leqq g$ mean that $f(y) \leqq g(y)$ for all y.

The reader will have no difficulty in verifying that $X + Y$, $X \cdot Y$ and X^Y are partially ordered systems. Further, if X and Y are *lattices*,[4] then so is $X \cdot Y$. Moreover any laws such as the modular and distributive laws which hold in X and Y hold in $X \cdot Y$. While if X is a lattice and n is the power of Y, then X^Y is a sublattice of X^n, and so X^Y is a modular resp. distributive lattice if X is.

The proofs of these facts will be omitted. Also, we shall not prove that the above definitions yield an extension of the cardinal arithmetic of Hausdorff—this is evident if one looks at Hausdorff's definitions (loc. cit.).

3. **Applications.** It is interesting to consider various arithmetic combinations of especially simple partially ordered systems, which have an independent algebraic importance.

As regards the simple systems, we shall let n denote the unordered aggregate of n elements, C_n the sequence of n elements (n finite), symbols $\aleph_\alpha$ the transfinite cardinals (= unordered aggregates), and adopt the conventional notation for transfinite ordinals. Finally, we shall let $B = C_2$ denote the Boolean algebra of two elements, and P_n the one-dimensional projective geometry with n points on its line. Then

(3.1) The finite Boolean algebras are the B^n.

(3.2) The Boolean algebra of all subsets of any aggregate of power $\aleph$ is $B^\aleph$.

(3.3) The finite distributive lattices are the B^X, where X denotes a variable finite partially ordered set.

(3.4) The "quotient-lattice" associated by Ore[5] with each abstract lattice L is L^B.

[4] By a "lattice" is meant a partially ordered system L in which any two elements x and x' have a g.l.b. $x \cap x'$ such that $x'' \leqq x \cap x'$ means that $x'' \leqq x$ and $x'' \leqq x'$, and a l.u.b. $x \cup x'$ such that $x'' \geqq x \cup x'$ means that $x'' \geqq x$ and $x'' \geqq x'$. By the *modular* law is meant the law that $x \leqq x''$ implies $x \cup (x' \cap x'') = (x \cup x') \cap x''$; the *distributive* law asserts that $(x \cup x') \cap (x' \cup x'') \cap (x'' \cup x) = (x \cap x') \cup (x' \cap x'') \cup (x'' \cap x)$.

[5] *On the foundations of abstract algebra.* I, Annals of Math., vol. 36 (1935), p. 425. Ore calls lattices *structures*.

(3.5) The integers ordered with respect to divisibility (the relation $m \mid n$) are a sublattice of $\omega^{\aleph_0}$.
(3.6) The "free" Boolean algebra generated by n symbols is B^{2^n}.
(3.7) The "free" distributive lattice generated by n symbols and with 0 and I added, is B^{B^n}.
(3.8) The "free" modular lattice generated by three symbols is a sublattice of $B^3 \cdot P_3$.
(3.9) The most general configuration generated in n-space by an r-plane and an s-plane through the origin, after iterated sections, linear sums and taking of orthogonal complements, is $B^4 \cdot P_4$.

The proofs of (3.1)–(3.9) are various. Assertions (3.1), (3.2) and (3.6) are known.[6] Assertions (3.3) and (3.7) are proved in the author's article, *Rings of sets*, which will appear in the next issue of this journal. Inspection of Ore's definition, which is identical with our definition of L^B, yields (3.4). Statement (3.5) is a corollary of the unique representation of any integer as a product of powers of ascending primes. The result (3.8) has been proved by the author (*On the structure of abstract algebras*, Proc. Camb. Phil. Soc., vol. 31 (1935), p. 443, Theorem 14), while (3.9) has been recently proved by J. von Neumann; the proof will be published elsewhere.

4. **Arithmetic identities.** It is obvious that addition and multiplication are commutative—in symbols, that

$$X + Y = Y + X \quad \text{and} \quad X \cdot Y = Y \cdot X. \tag{4.1}$$

They are also associative—that is,

$$X + (Y + Z) = (X + Y) + Z \quad \text{and} \quad X \cdot (Y \cdot Z) = (X \cdot Y) \cdot Z. \tag{4.2}$$

For both $(X + Y) + Z$ and $X + (Y + Z)$ consist of all members of either X, Y, or Z, with the provision that $x \leqq x'$, $y \leqq y'$, and $z \leqq z'$ preserve their meaning, while $x \leqq y$, $x \geqq y$, $y \leqq z$, $y \geqq z$, $z \leqq x$, and $z \geqq x$ are always denied. And both $X \cdot (Y \cdot Z)$ and $(X \cdot Y) \cdot Z$ consist of all triples $[x, y, z]$ with $x \in X$, $y \in Y$ and $z \in Z$, where $[x, y, z] \leqq [x', y', z']$ means that $x \leqq x'$, $y \leqq x'$ and $z \leqq z'$. Again

$$X \cdot (Y + Z) = X \cdot Y + X \cdot Z \quad \text{and} \quad (X + Y) \cdot Z = X \cdot Z + Y \cdot Z. \tag{4.3}$$

(In words, multiplication is distributive with respect to addition.) For $X \cdot (Y + Z)$ and $X \cdot Y + X \cdot Z$ alike consist of all couples $[x, y]$ and $[x, z]$ with $x \in X$, $y \in Y$, and $z \in Z$, where $[x, y] \leqq [x', y']$ means that $x \leqq x'$ and $y \leqq y'$, $[x, z] \leqq [x', z']$ means that $x \leqq x'$ and $z \leqq z'$, while $[x, y] \leqq [x', z]$ and $[x, y] \geqq [x', z]$ are always denied. This proves right-distributivity; left-distributivity follows by commutativity.

[6] For instance (3.1) is proved as Theorem 25.1 of the author's *On the combination of subalgebras*, Proc. Camb. Phil. Soc., vol. 29 (1933), p. 460; (3.6) follows from this and E. Schröder's *Algebra der Logik*; (3.2) is immediately obvious.

Again, although exponentiation is not commutative—in general, $X^Y \neq Y^X$—it satisfies the usual identities

$$(4.4)\qquad X^Y X^Z = X^{Y+Z} \quad\text{and}\quad X^Z Y^Z = (XY)^Z,$$

$$(4.5)\qquad X^1 = X \quad\text{and}\quad (X^Y)^Z = X^{YZ}.$$

The first identity of (4.4) is easy to prove. The system $X^Y X^Z$ consists of all couples $[f, g]$ where f and g have Y resp. Z for domains, and so are equivalent (Y and Z being non-overlapping in $Y + Z$) to a single function h with domain $Y + Z$. Moreover $[f, g] \leqq [f', g']$ means that for all $y \in Y$ and $z \in Z$, $f(y) \leqq f'(y)$ and $g(z) \leqq g'(z)$—that is, that for all $u \in (Y + Z)$, $h(u) \leqq h'(u)$. Hence $X^Y \cdot X^Z = X^{Y+Z}$.

Again, $X^Z Y^Z$ is the set of all function-couples $[f, g]$, where f and g are from Z to X resp. Z to Y, and $[f, g] = [f', g']$ means that $f(z) \leqq f'(z)$ and $g(z) \leqq g'(z)$ for all z. But each such $[f, g]$ can be regarded as a function h carrying each $z \in Z$ into $[f(z), g(z)] \in XY$—and so, since $h \leqq h'$ if and only if $[f, g] \leqq [f', g']$, $X^Z Y^Z = (XY)^Z$.

That $X^1 = X$ is obvious; it is a corollary (using (4.4)) that $X^2 = XX$, $X^3 = XXX, \cdots$, and that $X^m X^n = X^{m+n}$.

Actually, the second half of (4.5) is not simple to prove, and we shall start by analyzing X^{YZ}. By definition, this consists of all monotonic functions f with arguments $[y, z]$ and values x $[x \in X, y \in Y, z \in Z]$. Such functions f associate with each $z \in Z$, a function g_z with arguments $y \in Y$ and values $g_z(y) = f([y, z])$ in X. But for any z, $y \leqq y'$ implies $[y, z] \leqq [y', z]$ in YZ, hence

$$g_z(y) = f([y, z]) \leqq f([y', z]) = g_z(y')$$

and so g_z is in X^Y.

Furthermore, if $z \leqq z'$ and y is fixed, then

$$g_z(y) = f([y, z]) \leqq f([y, z']) = g_{z'}(y),$$

and so the correspondence $f\colon z \to g_z$ is monotonic, which makes X^{YZ} a subset S of $(X^Y)^Z$. While $f \leqq f'$ if and only if $f([y, z]) \leqq f'([y, z])$ for all $[y, z]$—which means that for all z, $g_z(y) \leqq g_z'(y)$ for all y, which means that for all z, $g_z \leqq g_z'$, and so X^{YZ} is isomorphic with S. But conversely, given $f \in (X^Y)^Z$, if $y \leqq y'$ and $z \leqq z'$, then $g_z(y) \leqq g_{z'}(y) \leqq g_{z'}(y')$, and so $f \in X^{YZ}$, which completes the proof of the fact that X^{YZ} and $(X^Y)^Z$ are isomorphic.

5. **Decomposition theorems.** It is evident that if $X = X_1 + \cdots + X_r = Y_1 + \cdots + Y_s$ is any partially ordered system, written in two ways as the sum of additive components, then by set-theory $X = Z_{11} + \cdots + Z_{rs}$, where $Z_{ij} = X_i \cap Y_j$ denotes the set of elements common to X_i and Y_j, ordered as in X. For if $x \in Z_{ij}$ and $x \leqq x'$, then $x' \in X_i$ and $x' \in Y_j$, whence $x' \in Z_{ij}$.

It follows that if X is finite, it has one and only one representation $X = X_1 + \cdots + X_q$ as the sum of (non-void) partially ordered sets not themselves further reducible. Actually, if we regard the elements of X as vertices of a

graph, and join two vertices x and x' when $x < x'$ or $x > x'$ and no element z with $x < z < x'$ or $x > z > x'$ can be interpolated between x and x', then the irreducible additive components of X are the connected components of its graph.

The author has proved elsewhere[7] that any finite *lattice* L has a similar unique representation $L = L_1 \cdots L_q$ as the product of constituents not themselves further reducible. The prime-factor theorem of arithmetic asserts that the same conclusion is valid if L is a totally unordered aggregate.

Can we unite both these results in the single assertion that each finite *partially ordered system* X can be written in one and only one way as a product $X_1 \cdots X_q$ of "prime" (= indecomposable) factors? It is quite possible.

6. **Solution of equations.** One naturally asks if it is true that in the finite case, the rules that

$$X + Y = X + Z \quad \text{implies} \quad Y = Z, \tag{6.1}$$

$$X \cdot Y = X \cdot Z \quad \text{implies} \quad Y = Z, \tag{6.2}$$

which are true for pure cardinals and ordinals, are valid for general partially ordered sets.

It is easy to prove (6.1) by appealing to the result that every finite partially ordered system X has a unique decomposition into a sum of irreducible components. For, making this decomposition, we have

$$X + Y = (X_1 + \cdots + X_q) + (Y_1 + \cdots + Y_r),$$

$$X + Z = (X_1 + \cdots + X_q) + (Z_1 + \cdots + Z_s).$$

One sees that by ordinary subtraction, each kind of component must occur in Y and in Z alike, the number of times it occurs in $X + Y = X + Z$ minus the number of times it occurs in X, whence $Y = Z$.

This argument would prove (6.2) similarly if we knew that the proposition stated at the end of §5 was true; in any event, it holds for lattices, since they have unique decompositions into irreducible multiplicative factors.

If it held in general, would it imply that one could, by introducing negatives and quotients as ideal elements, extend our arithmetic to form an abstract field of elements of the symbolic form $(X - Y)/Z$? This seems highly improbable.

7. **Lexicographic combinations.** Let X be any partially ordered system whose members are themselves (not necessarily distinct) partially ordered systems Y_x.

We shall define $\sum_X Y_x$ (in words, the lexicographic sum of the Y_x relative to the ordering X) as the system whose members are the $y_x \in Y_x$ $[x \in X]$, in which $y_x \leq y'_x$ preserves its meaning in Y_x, and $y_x \leq y'_{x'}$ $[x \neq x']$ means that $x < x'$.

[7] *On the lattice theory of ideals*, Bull. Am. Math. Soc., vol. 40 (1934), p. 616, Theorem 2 and its corollaries.

Similarly, we shall define $\prod_X Y_x$ (in words, the lexicographic product of the Y_x relative to the ordering X) as the system whose members are the functions f carrying each $x \in X$ into an $f(x) \in Y_x$, and in which $f \leqq g$ means that for every x_0, either $f(x_0) \leqq g(x_0)$ or there exists an $\bar{x}_0 < x_0$ such that $f(\bar{x}_0) < g(\bar{x}_0)$ and $f(x) \leqq g(x)$ for all $x < \bar{x}_0$.

In case X is totally unordered $\sum_X Y_x$ and $\prod_X Y_x$ are the *cardinal* sums and products of Hausdorff extended as in §3. In case the Y_x are sequences arranged in the sequential order X, $\sum_X Y_x$ and $\prod_X Y_x$ are the *ordinal* sums and products of Hausdorff (loc. cit.).

If all the Y_x are isomorphic—that is, letting them all be a particular partially ordered system Y—we get a new binary operation of exponential type. For although the lexicographic sum $\sum_X Y$ of the occurrences of Y relative to the ordering X is simply the lexicographic product $\prod_{X<Y} X \cdot Y$ of X and Y relative to the ordering $X < Y$, the lexicographic product $\prod_X Y$, which we may denote lex Y^X, is a new system which could not easily be defined otherwise. Its elements are the different "words" f with letters from X, one in each position of Y, and ordered in lexicographic order.

Lexicographic combination is non-commutative, and seems to have no interest aside from the fact that lexicographic combinations of simply ordered (or well-ordered) systems are always themselves simply ordered (resp. well-ordered). This property makes them useful in ordering words in dictionaries.

HARVARD UNIVERSITY.

THE IMBEDDING OF RIEMANN SPACES IN THE LARGE

By Carl B. Allendoerfer

1. The "classical" theory which describes the properties of an n-dimensional Riemann space immersed in an $(n + p)$-dimensional Euclidean space has its roots in the work of Voss and Ricci,[1] and is a natural generalization of the properties of a two-dimensional surface in ordinary three-space.

A more recent theory is that of W. Mayer, the latest refinement of which appeared in the Transactions of the American Mathematical Society in 1935 (vol. 38, pp. 267–309). We shall refer to this paper as "M". This theory combines the generalization of the properties of an ordinary surface with the generalization of those of a curve in an n-dimensional Euclidean space. Thus the theory speaks of the "first normal space" of an n-dimensional Riemann space as an analogue of the first ("principal") normal of a curve. Second and higher normal spaces also occur as extensions of the notions of the second and higher order normals of a curve. Associated with each normal space is a fundamental form, the form for the r-th normal space being of the $2(r + 1)$-th degree. These forms give a complete description of the geometry of the normal spaces. Mayer has shown that the coefficients of these forms are not entirely independent, but satisfy certain relations. The first object of this paper is to prove some additional relations of this nature and to determine a set of them which is the necessary and sufficient condition on a number of forms that they describe a Riemann space which is actually imbedded in a Euclidean space. These results are stated in Theorem VII. In the course of the argument (§5) we prove a new set of identities in the curvature tensors of order higher than the first which are analogous to those for the classical Riemann tensor.

The previous treatment of Mayer's theory had the further defect that it was purely local, referring to an unspecified domain about a suitable point. In this paper the differential equations which occur are investigated with the purpose of finding the maximum region within which they may be integrated. We define singular points and show that the results of Theorem VII hold for any simply connected portion of the space which does not contain a singular point.

Received December 1, 1936.

[1] A. Voss, *Zur Theorie der Transformation quadratischer Differentialausdrücke und der Krümmung höherer Mannigfaltigkeiten*, Math. Ann., vol. 16 (1880), pp. 129–179.

G. Ricci, *Formole fondamentali nella teoria generale di varietà e della loro curvatura*, Rend. dei Lincei, vol. 11 (1902), pp. 355–362.

A modern exposition of the theory is given by Eisenhart, *Riemannian Geometry*, 1925, Chapter IV.

This treatment is in the spirit of recent work by T. Y. Thomas and W. Mayer[2] in similar connections, and many of the methods used here are due to those writers.

2. **Riemann manifolds.** We shall consider a Riemann manifold, R_n, i.e., a coördinate manifold endowed with the Riemann metric

$$ds^2 = g_{\alpha\beta} dx^\alpha dx^\beta, \tag{2.1}$$

where the $g_{\alpha\beta}(x)$ are continuous functions having continuous derivatives to a definite order. This order will be specified in the next section. We assume that the determinant $| g_{\alpha\beta} |$ is positive definite. Let y^i $(i = 1, \cdots, n + p)$ be the rectangular Cartesian coördinates of an $(n + p)$-dimensional Euclidean space E_{n+p}. Then R_n is said to be imbedded in E_{n+p} if the equations

$$g_{\alpha\beta} = \frac{\partial y^i}{\partial x^\alpha} \frac{\partial y^i}{\partial x^\beta} \tag{2.2}$$

have a solution

$$y^i = \varphi^i(x) \tag{2.3}$$

in each coördinate neighborhood of R_n. We shall understand hereafter that the coördinate system x is a typical coördinate system, not some special one, and consequently an invariant equation such as (2.2) will be understood to hold in all coördinate neighborhoods without further mention of this fact.

3. **The normal spaces.** For each value of α and for linear transformations of the y's, the quantities $\dfrac{\partial y^i}{\partial x^\alpha}$ are the components of a vector field defined for all values of $x \subset U$. The totality of vector fields defined by $\sigma^\alpha \dfrac{\partial y^i}{\partial x^\alpha}$ (where σ^α are arbitrary scalars in E_{n+p}) is called the first osculating space I_1. In a similar manner, the set of all vector fields defined by

$$\sum\left(\sigma^{\alpha_1} \frac{\partial y^i}{\partial x^{\alpha_1}} + \sigma^{\alpha_1\alpha_2} \frac{\partial^2 y^i}{\partial x^{\alpha_1} \partial x^{\alpha_2}} + \cdots + \sigma^{\alpha_1\alpha_2\cdots\alpha_h} \frac{\partial^h y^i}{\partial x^{\alpha_1} \cdots \partial x^{\alpha_h}}\right) \tag{3.1}$$

(where again the σ's are scalars) is called the h-th osculating space $I_{12\cdots h}$. The invariance of these spaces under transformations of the x's is proved in M, p. 270.

The totality of vector fields belonging to $I_{12\cdots h}$ and normal to $I_{12\cdots h-1}$ is called the $(h - 1)$-th normal space, I_h. We shall denote the projection of the vector field

$$\frac{\partial^h y^i}{\partial x^{\alpha_1} \cdots \partial x^{\alpha_h}}$$

[2] T. Y. Thomas, *Riemann spaces of class one and their characterization*, Acta Mathematica, vol. 67 (1936), pp. 169–211; *Fields of parallel vectors in the large*, Compositio Mathematica, vol. 3 (1936), pp. 453–468.

into the normal space I_h by $Y^i_{\alpha_1 \cdots \alpha_h}$. Since the vector fields belonging to each normal space are normal to those belonging to every other normal space, we have

$$Y^i_{\alpha_1 \cdots \alpha_h} Y^i_{\beta_1 \cdots \beta_k} = 0 \qquad (h \neq k). \tag{3.2}$$

Since there are at most p independent normal vectors, it is clear that there are at most $(m - 1) \leqq p$ normal spaces. It is therefore of significance to speak of the "last" normal space, I_m. The index m will have this significance throughout the paper.

Let us now assume that the functions $\varphi^i(x)$ occurring in (2.3) are of class C^{m+1}, that is, they have continuous derivatives of the $(m + 1)$-th order. The coefficients of the first fundamental form $g_{\alpha\beta}$ given by (2.4) are therefore of class C^m. We shall also denote them by $E_{\alpha\beta}(\equiv g_{\alpha\beta})$. The coefficients of the h-th fundamental form are defined by

$$E_{\alpha_1 \cdots \alpha_h | \beta_1 \cdots \beta_h} = Y^i_{\alpha_1 \cdots \alpha_h} Y^i_{\beta_1 \cdots \beta_h} \tag{3.3}$$

for $h = 1, \cdots, m$. From their definitions it is clear that $E_{\alpha_1 \cdots \alpha_h | \beta_1 \cdots \beta_h}$ are of class C^{m-h+1}. They are also tensors under transformations of the x's, for it follows from their law of transformation that $Y^i_{\alpha_1 \cdots \alpha_h}$ are tensors. Because of (3.3) the matrix[3]

$$E_h \equiv \| E_{\alpha_1 \cdots \alpha_h | \beta_1 \cdots \beta_h} \|$$

is positive semi-definite and its rank is the same as that of the matrix

$$Y_h \equiv \| Y^i_{\alpha_1 \cdots \alpha_h} \|,$$

where i denotes the rows and combinations of $\alpha_1 \cdots \alpha_h$ denote the columns. In fact (cf. M, p. 274) the equations

$$E_{\alpha_1 \cdots \alpha_h | \beta_1 \cdots \beta_h} \theta^{\beta_1 \cdots \beta_h} = 0 \tag{3.4}$$

and

$$Y^i_{\beta_1 \cdots \beta_h} \theta^{\beta_1 \cdots \beta_h} = 0 \tag{3.5}$$

have the same solutions at any point P.

A point $P \subset R_n$ will be called a regular point of the imbedding if there exists a neighborhood $N(P) \supset P$ within which each Y_h is of constant rank. This implies that each E_h is of constant rank in $N(P)$, and conversely. A point which is not regular will be called singular. Consider now the matrix

$$E \equiv \left\| \begin{matrix} \boxed{E_{\alpha\beta}} & & & \\ & \boxed{E_{\alpha\beta|\gamma\delta}} & & \\ & & \ddots & \\ & & & \boxed{E_{\alpha_1 \cdots \alpha_m | \beta_1 \cdots \beta_m}} \end{matrix} \right\|.$$

[3] The rows are given by combinations of the indices $\alpha_1 \cdots \alpha_h$, and the columns are given by combinations of $\beta_1 \cdots \beta_h$.

It is clear that E is of constant rank in $N(P)$. Suppose, conversely, that in a neighborhood $M(P)$ the matrix E is of constant rank. Now there exists a neighborhood $\bar{M}(P) \subset M(P)$ within which each E_h has a rank not lower than its rank at P. But since the rank of E is the sum of the ranks of E_h, and since the rank of E is constant in $\bar{M}(P)$, it follows that each E_h is of constant rank in $\bar{M}(P)$, i.e., P is a regular point. Hence we may make the equivalent

DEFINITION. *A point $P \subset R_n$ is regular if there exists a neighborhood $N(P)$ within which the matrix E is of constant rank. Other points are called singular points.*

It will become evident in the next section that the differential equations we consider leave the rank of E constant in any domain in which they can be integrated. Hence we see that our singular points are also singularities of the system of differential equations. We shall show them to be the only such points.

It is evident at once that the set of singular points is nowhere dense in R_n and is a variety of dimension lower than n. Hence there will exist a neighborhood of every regular point which contains only regular points, and in fact this neighborhood can be taken to be simply connected. *We shall call such a domain the neighborhood U*, and shall limit our discussion to regions of this nature.

4. **The Frenet equations.** It can be proved (M, p. 278) that at any point $P \subset U$ there exist Γ's which satisfy the Frenet equations

$$
(4.1)\qquad
\begin{aligned}
&\frac{\partial}{\partial x^{\alpha_1}}(y^i) = Y^i_{\alpha_1},\\
&\frac{\partial}{\partial x^{\alpha_2}}(Y^i_{\alpha_1}) = \Gamma^{\beta_1}_{\alpha_1\alpha_2} Y^i_{\beta_1} + Y^i_{\alpha_1\alpha_2},\\
&\cdots\cdots\cdots\cdots\\
&\frac{\partial}{\partial x^{\alpha_{h+1}}}(Y^i_{\alpha_1\cdots\alpha_h}) = \Gamma^{\beta_1\cdots\beta_{h-1}}_{\alpha_1\cdots\alpha_{h+1}} Y^i_{\beta_1\cdots\beta_{h-1}} + \Gamma^{\beta_1\cdots\beta_h}_{\alpha_1\cdots\alpha_{h+1}} Y^i_{\beta_1\cdots\beta_h} + Y^i_{\alpha_1\cdots\alpha_{h+1}},\\
&\cdots\cdots\cdots\cdots\\
&\frac{\partial}{\partial x^{\alpha_{m+1}}}(Y^i_{\alpha_1\cdots\alpha_m}) = \Gamma^{\beta_1\cdots\beta_{m-1}}_{\alpha_1\cdots\alpha_{m+1}} Y^i_{\beta_1\cdots\beta_{m-1}} + \Gamma^{\beta_1\cdots\beta_m}_{\alpha_1\cdots\alpha_{m+1}} Y^i_{\beta_1\cdots\beta_m}.
\end{aligned}
$$

From (3.2), (3.3) and (4.1) it follows that for every point we have

$$
(4.2)\qquad E_{\beta_1\cdots\beta_h|\gamma_1\cdots\gamma_h}\Gamma^{\beta_1\cdots\beta_h}_{\alpha_1\cdots\alpha_h\tau} = \left[\frac{\partial}{\partial x^\tau}(Y^i_{\alpha_1\cdots\alpha_h})\right] Y^i_{\gamma_1\cdots\gamma_h}
$$

and

$$
(4.3)\qquad E_{\beta_1\cdots\beta_{h-1}|\gamma_1\cdots\gamma_{h-1}}\Gamma^{\beta_1\cdots\beta_{h-1}}_{\alpha_1\cdots\alpha_h\tau} = \left[\frac{\partial}{\partial x^\tau}(Y^i_{\alpha_1\cdots\alpha_h})\right] Y^i_{\gamma_1\cdots\gamma_{h-1}}.
$$

On account of (3.2) we may write (4.3) in the form

$$
(4.4)\qquad E_{\beta_1\cdots\beta_{h-1}|\gamma_1\cdots\gamma_{h-1}}\Gamma^{\beta_1\cdots\beta_{h-1}}_{\alpha_1\cdots\alpha_h\tau} = -Y^i_{\alpha_1\cdots\alpha_h}\frac{\partial}{\partial x^\tau}(Y^i_{\gamma_1\cdots\gamma_{h-1}}).
$$

Since by hypothesis the right hand sides of (4.2) and of (4.4) considered as functions of the x's are of class C^{m-h} and C^{m-h+1} respectively, we have the fact that the left sides have this same property.

Since we have assumed that each matrix E_h is of constant rank r_h in U, to every point $P \subset U$ there corresponds at least one non-vanishing r_h-th order minor, D_h, of each matrix E_h. Because of the continuity of the E's there exists a neighborhood $V(P) \supset P$ which is contained in U and within which no D_h vanishes.

Now we know that the Γ's actually exist, and so it is possible to solve (4.2) and (4.4) for the Γ's as functions of the x's within each $V(P)$. The solutions in general will be determined to within additive functions $\theta^{\beta_1 \cdots \beta_h}$ and $\theta^{\beta_1 \cdots \beta_{h-1}}$, respectively, which are solutions of

$$(4.5) \qquad E_{\beta_1 \cdots \beta_h | \gamma_1 \cdots \gamma_h}\, \theta^{\beta_1 \cdots \beta_h} = 0, \qquad E_{\beta_1 \cdots \beta_{h-1} | \gamma_1 \cdots \gamma_{h-1}}\, \theta^{\beta_1 \cdots \beta_{h-1}} = 0.$$

The θ's may be chosen so that $\Gamma^{\beta_1 \cdots \beta_h}_{\alpha_1 \cdots \alpha_h \tau}$ and $\Gamma^{\beta_1 \cdots \beta_{h-1}}_{\alpha_1 \cdots \alpha_h \tau}$ are of class C^{m-h} and C^{m-h+1} respectively in $V(P)$.

If $\Gamma^{\beta_1 \cdots \beta_h}_{\alpha_1 \cdots \alpha_h \tau}$ are solutions of (4.2) in $V(P)$ and if $\Gamma'^{\beta_1 \cdots \beta_h}_{\alpha_1 \cdots \alpha_h \tau}$ are solutions in another neighborhood $V(P')$, it is clear that in $V(P) \cap V(P')$

$$(\Gamma^{\beta_1 \cdots \beta_h}_{\alpha_1 \cdots \alpha_h \tau} - \Gamma'^{\beta_1 \cdots \beta_h}_{\alpha_1 \cdots \alpha_h \tau}) E_{\beta_1 \cdots \beta_h | \gamma_1 \cdots \gamma_h} = 0.$$

A similar result holds for $\Gamma'^{\beta_1 \cdots \beta_{h-1}}_{\alpha_1 \cdots \alpha_h \tau}$. Mayer has proved (M, p. 279) that for transformations of the x's, $\Gamma^{\beta_1 \cdots \beta_{h-1}}_{\alpha_1 \cdots \alpha_h \tau}$ are tensors, and that $\Gamma^{\beta_1 \cdots \beta_h}_{\alpha_1 \cdots \alpha_h \tau}$ have a transformation law like that of the Christoffel symbols.

If we form the integrability conditions of (4.1) in any $V(P)$, we obtain

$$(4.6) \qquad (\Gamma^{\rho_1 \cdots \rho_{h-1}}_{\alpha_1 \cdots \alpha_h \tau}\, \Gamma^{\sigma_1 \cdots \sigma_{h-2}}_{\rho_1 \cdots \rho_{h-1} \omega} - \Gamma^{\rho_1 \cdots \rho_{h-1}}_{\alpha_1 \cdots \alpha_h \omega}\, \Gamma^{\sigma_1 \cdots \sigma_{h-2}}_{\rho_1 \cdots \rho_{h-1} \tau})\, Y^i_{\sigma_1 \cdots \sigma_{h-2}} = 0;$$

$$(4.7) \qquad \begin{aligned} \Bigg(&\frac{\partial \Gamma^{\sigma_1 \cdots \sigma_{h-1}}_{\alpha_1 \cdots \alpha_h \tau}}{\partial x^\omega} - \frac{\partial \Gamma^{\sigma_1 \cdots \sigma_{h-1}}_{\alpha_1 \cdots \alpha_h \omega}}{\partial x^\tau} + \Gamma^{\rho_1 \cdots \rho_{h-1}}_{\alpha_1 \cdots \alpha_h \tau}\, \Gamma^{\sigma_1 \cdots \sigma_{h-1}}_{\rho_1 \cdots \rho_{h-1} \omega} - \Gamma^{\rho_1 \cdots \rho_{h-1}}_{\alpha_1 \cdots \alpha_h \omega}\, \Gamma^{\sigma_1 \cdots \sigma_{h-1}}_{\rho_1 \cdots \rho_{h-1}} \\ &+ \Gamma^{\rho_1 \cdots \rho_h}_{\alpha_1 \cdots \alpha_h \tau}\, \Gamma^{\sigma_1 \cdots \sigma_{h-1}}_{\rho_1 \cdots \rho_h \omega} - \Gamma^{\rho_1 \cdots \rho_h}_{\alpha_1 \cdots \alpha_h \omega}\, \Gamma^{\sigma_1 \cdots \sigma_{h-1}}_{\rho_1 \cdots \rho_h \tau} \Bigg) Y^i_{\sigma_1 \cdots \sigma_{h-1}} = 0; \end{aligned}$$

$$(4.8) \qquad \begin{aligned} \Bigg(&\frac{\partial \Gamma^{\sigma_1 \cdots \sigma_h}_{\alpha_1 \cdots \alpha_h \tau}}{\partial x^\omega} - \frac{\partial \Gamma^{\sigma_1 \cdots \sigma_h}_{\alpha_1 \cdots \alpha_h \omega}}{\partial x^\tau} + \Gamma^{\rho_1 \cdots \rho_h}_{\alpha_1 \cdots \alpha_h \tau}\, \Gamma^{\sigma_1 \cdots \sigma_h}_{\rho_1 \cdots \rho_h \omega} - \Gamma^{\rho_1 \cdots \rho_h}_{\alpha_1 \cdots \alpha_h \omega}\, \Gamma^{\sigma_1 \cdots \sigma_h}_{\rho_1 \cdots \rho_h \tau} \\ &+ \Gamma^{\sigma_1 \cdots \sigma_{h-1}}_{\alpha_1 \cdots \alpha_h \tau}\, \delta^{\sigma_h}_\omega - \Gamma^{\sigma_1 \cdots \sigma_{h-1}}_{\alpha_1 \cdots \alpha_h \omega}\, \delta^{\sigma_h}_\tau + \Gamma^{\sigma_1 \cdots \sigma_h}_{\alpha_1 \cdots \alpha_h \tau \omega} - \Gamma^{\sigma_1 \cdots \sigma_h}_{\alpha_1 \cdots \alpha_h \omega \tau} \Bigg) Y^i_{\sigma_1 \cdots \sigma_h} = 0, \end{aligned}$$

or

$$(B^{\sigma_1 \cdots \sigma_h}_{\alpha_1 \cdots \alpha_h \tau \omega} + \Gamma^{\sigma_1 \cdots \sigma_h}_{\alpha_1 \cdots \alpha_h \tau \omega} - \Gamma^{\sigma_1 \cdots \sigma_h}_{\alpha_1 \cdots \alpha_h \omega \tau}) Y^i_{\sigma_1 \cdots \sigma_h} = 0,$$

where B's are thus defined. When $h = m$ we have

$$(4.9) \qquad B^{\sigma_1 \cdots \sigma_m}_{\alpha_1 \cdots \alpha_m \tau \omega}\, Y^i_{\sigma_1 \cdots \sigma_m} = 0.$$

Finally we have

$$(4.10) \quad (\Gamma^{\sigma_1 \cdots \sigma_{h-1}}_{\alpha_1 \cdots \alpha_{h-1} \tau}\, \delta^{\sigma_h}_\omega - \Gamma^{\sigma_1 \cdots \sigma_{h-1}}_{\alpha_1 \cdots \alpha_{h-1} \omega}\, \delta^{\sigma_h}_\tau + \Gamma^{\sigma_1 \cdots \sigma_h}_{\alpha_1 \cdots \alpha_{h-1} \tau \omega} - \Gamma^{\sigma_1 \cdots \sigma_h}_{\alpha_1 \cdots \alpha_{h-1} \omega \tau}) Y^i_{\sigma_1 \cdots \sigma_h} = 0.$$

In these equations the index h has the range $1, \cdots, m$, it being understood that no terms involving zero or negative indices are written. The entire set may be written in a more convenient notation by writing s_h for the set of h symmetric indices $\sigma_1 \cdots \sigma_h$. A single latin index without a subscript stands for a single greek index. Thus, for example, we write $\Gamma^{\sigma_1 \cdots \sigma_h}_{\alpha_1 \cdots \alpha_h \tau} \equiv \Gamma^{s_h}_{a_h t}$. Hereafter we shall pass from one notation to the other without comment. If (4.6) to (4.10) be multiplied by appropriate Y's and the summations performed (for example multiply (4.6) by $Y^i_{\beta_1 \cdots \beta_{h-2}}$), the set can be written in the shorter notation as follows:

$$(4.11) \qquad (\Gamma^{r_{h-1}}_{a_h t} \Gamma^{s_{h-2}}_{r_{h-1} w} - \Gamma^{r_{h-1}}_{a_h w} \Gamma^{s_{h-2}}_{r_{h-1} t}) E_{s_{h-2} | b_{h-2}} = 0;$$

$$(4.12) \qquad \begin{aligned} \biggl(\frac{\partial \Gamma^{s_{h-1}}_{a_h t}}{\partial x^w} - \frac{\partial \Gamma^{s_{h-1}}_{a_h w}}{\partial x^t} &+ \Gamma^{r_{h-1}}_{a_h t} \Gamma^{s_{h-1}}_{r_{h-1} w} - \Gamma^{r_{h-1}}_{a_h w} \Gamma^{s_{h-1}}_{r_{h-1} t} \\ &+ \Gamma^{r_h}_{a_h t} \Gamma^{s_{h-1}}_{r_h w} - \Gamma^{r_h}_{a_h w} \Gamma^{s_{h-1}}_{r_h t} \biggr) E_{s_{h-1} | b_{h-1}} = 0; \end{aligned}$$

$$(4.13) \qquad (B^{s_h}_{a_h t w} + \Gamma^{s_h}_{a_h t w} - \Gamma^{s_h}_{a_h w t}) E_{s_h | b_h} = 0 \qquad (h = 1, \cdots, m-1);$$

$$(4.14) \qquad B^{s_m}_{a_m t w} E_{s_m | b_m} = 0;$$

$$(4.15) \qquad (\Gamma^{s_{h-1}}_{a_{h-1} t} \delta^r_w - \Gamma^{s_{h-1}}_{a_{h-1} w} \delta^r_t + \Gamma^{s_{h-1} r}_{a_{h-1} t w} - \Gamma^{s_{h-1} r}_{a_{h-1} w t}) E_{s_{h-1} r | b_h} = 0.$$

If we differentiate (3.2) and (3.3), substitute from (4.1), and again make use of (3.2) and (3.3), we have for $h = 1, \cdots, m$

$$(4.16) \qquad E_{a_{h-1} t | b_h} + \Gamma^{d_{h-1}}_{b_h t} E_{d_{h-1} | a_{h-1}} = 0;$$

$$(4.17) \qquad \frac{\partial}{\partial x^t} (E_{a_h | b_h}) - \Gamma^{d_h}_{a_h t} E_{d_h | b_h} - \Gamma^{d_h}_{b_h t} E_{d_h | a_h} = 0.$$

The entire set (4.11) to (4.17) hold in every $V(P)$.

THEOREM I. *If $y^i = \varphi^i(x)$ are functions of class C^{m+1} which define a Riemann space imbedded in a given E_{n+p} throughout a neighborhood U, then*

1. *The tensors $E_{\alpha_1 \cdots \alpha_h | \beta_1 \cdots \beta_h}$ $(h = 1, \cdots, m)$ are defined as functions of class C^{m-h+1} in U; the matrix E_1 is positive definite in U, and the matrices E_h $(h = 2, \cdots, m)$ are positive semi-definite of constant rank r_h in U and $r_2 + \cdots + r_m = p$.*

2. *The quantities $\Gamma^{\beta_1 \cdots \beta_{h-1}}_{\alpha_1 \cdots \alpha_h \tau}(x)$ and $\Gamma^{\beta_1 \cdots \beta_h}_{\alpha_1 \cdots \alpha_h \tau}(x)$ are defined to within a solution of (4.5) in each $V(P) \subset U$ and are of class C^{m-h+1} and C^{m-h}, respectively.*

3. *The E's, the Γ's, and their derivatives satisfy* (4.11) *to* (4.17) *in each neighborhood $V(P)$.*

5. **Some identities and formal relations.** We interrupt the geometrical argument at this point to consider some formal relations which we shall need later. The expressions

$$B_{a_h | p_h | t w} \equiv B^{s_h}_{p_h t w} E_{s_h | a_h} \qquad (h = 1, \cdots, m)$$

are called the components of the h-th curvature tensor, the first curvature tensor being in fact the ordinary Riemann tensor. From (4.13) and (4.16) it follows that

$$B_{a_k|p_k|tw} = E_{a_kw|p_kt} - E_{a_kt|p_kw} \qquad (k = 1, \cdots, m-1). \tag{5.1}$$

These equations are generalizations of the Gauss equations of the classical theory, so we shall continue to call them the Gauss equations. Since the right sides of (5.1) are of class C^{m-k}, it follows that $B_{a_k|p_k|tw}$ are of class C^{m-k}. In fact, we see that the curvature tensors are defined as functions of this class throughout U in spite of the fact that the Γ's from which they are formed are defined only within each $V(P) \subset U$. It is also clear that the arbitrariness which occurred in the definition of the Γ's has disappeared from the B's. Because of (5.1) there exists a set of relations in the B's which holds for all $x \subset U$ and which is a generalization of that which holds for the Riemann tensor. By examining the definition of the B's, it can be shown that these relations are actually identities in the x's. They are the following:

$$B_{a_k|p_k|tw} + B_{a_k|p_k|wt} = 0; \tag{5.2}$$

$$B_{a_k|p_k|tw} + B_{p_k|a_k|tw} = 0; \tag{5.3}$$

$$B_{a_k|p_{k-1}v|tw} + B_{a_k|p_{k-1}w|vt} + B_{a_k|p_{k-1}t|wv} = 0; \tag{5.4}$$

$$P(B_{\alpha_1\cdots\alpha_k|\beta_1\cdots\beta_k|\tau\bar{\omega}}) = 0, \tag{5.5}$$

where P denotes the sum of terms formed by the cyclic interchange of $(\alpha_1 \cdots \alpha_k \beta_1 \cdots \beta_k\tau)$. We indicate by a bar over an index that it is not to be included in the permutation. The final identity is

$$S(B_{\alpha_1\cdots\alpha_k|\beta_1\cdots\beta_k|\tau\omega}) = 0, \tag{5.6}$$

where S indicates the sum of $k + 1$ terms, of which the first is the term written above, and the q-th of which is obtained from the $(q - 1)$-th by interchanging α_{q-1} with the last index of the $(q - 1)$-th term and β_{q-1} with the next to last index of the same term, and then interchanging α_{q-1} and β_{q-1} in the result. In case $k = 1$ this becomes the ordinary identity

$$B_{\alpha|\beta|\tau\omega} + B_{\omega|\tau|\alpha\beta} = 0.$$

For $k = 2$ we have

$$B_{\alpha_1\alpha_2|\beta_1\beta_2|\tau\omega} + B_{\omega\alpha_2|\tau\beta_2|\alpha_1\beta_1} + B_{\omega\beta_1|\tau\alpha_1|\alpha_2\beta_2} = 0.$$

In order to generalize the Bianchi identity we define a generalized covariant derivative which will be denoted by a semi-colon and which is a special case of those considered by A. W. Tucker.[4] The rule to be used is shown by the follow-

[4] A. W. Tucker, *On generalized covariant differentiation*, Annals of Math., vol. 32 (1931), pp. 451–460.

ing example, where t_f and r_h are compound indices of the same or different orders:

$$T^{t_f}_{r_h;w} = \frac{\partial T^{t_f}_{r_h}}{\partial x^w} + T^{q_f}_{r_h}\Gamma^{t_f}_{q_f w} - T^{t_f}_{p_h}\Gamma^{p_h}_{r_h w}. \tag{5.7}$$

The ordinary rules of covariant differentiation apply in this case, and for example we may write (4.17) in the form

$$E_{a_h|b_h;t} = 0.$$

When $h = 1$ and $f = 1$, we have ordinary covariant differentiation with respect to $\Gamma^{\alpha}_{\beta\gamma}$; *this will be denoted by a comma.* The generalized identity may be proved by direct calculation from (4.11)-(4.17) and has the form

$$B_{a_h|p_h|tw;v} + B_{a_h|p_h|vt;w} + B_{a_h|p_h|wv;t} = 0. \tag{5.8}$$

If (5.6) is differentiated covariantly and each term in the result is replaced by the corresponding two terms given by (5.8), we have

$$I(B_{\alpha_1\cdots\alpha_h|\beta_1\cdots\beta_h|\bar{\tau}\omega;\theta}) = 0, \tag{5.9}$$

where I denotes the sum of terms obtained by the cyclic interchange of

$$(\alpha_1 \cdots \alpha_h\,\theta\beta_1 \cdots \beta_h\,\omega).$$

By differentiating (5.5) we have

$$P(B_{\alpha_1\cdots\alpha_h|\beta_1\cdots\beta_h|\tau\bar{\omega};\bar{\lambda}}) = 0,$$

the cyclic interchange omitting ω and λ. Adding to this the identity

$$P(B_{\alpha_1\cdots\alpha_h|\beta_1\cdots\beta_h|\bar{\lambda}\tau;\bar{\omega}}) = 0,$$

and using (5.8), we have

$$P(B_{\alpha_1\cdots\alpha_h|\beta_1\cdots\beta_h|\bar{\omega}\bar{\lambda};\tau}) = 0. \tag{5.10}$$

This may be thought of as another generalization of Bianchi's identity.

It is also of importance to show that (4.12) is a consequence of the other equations of the set (4.11)-(4.17). Differentiating (4.16) and eliminating derivatives of the E's by (4.17), we have

$$E_{r_{h-1}|a_{h-1}}\frac{\partial\Gamma^{r_{h-1}}_{b_{h-1}wt}}{\partial x^v} = -\Gamma^{r_{h-1}}_{b_{h-1}w}(\Gamma^{s_{h-1}}_{r_{h-1}v}E_{s_{h-1}|a_{h-1}} + \Gamma^{s_{h-1}}_{a_{h-1}v}E_{s_{h-1}|r_{h-1}})$$
$$-\Gamma^{s_h}_{a_{h-1}tv}E_{s_h|b_{h-1}w} - \Gamma^{s_h}_{b_{h-1}wv}E_{s_h|a_{h-1}t}.$$

Substituting this into (4.12), we obtain an expression which vanishes identically because of (4.15) and (4.16). Thus (4.12) is a consequence of (4.15), (4.16), and (4.17).

A final relation is obtained by substituting in (4.11) from (4.16). The result is

$$\Gamma^{r_{h-1}}_{p_h t}E_{r_{h-1}|a_{h-2}w} - \Gamma^{r_{h-1}}_{p_h w}E_{r_{h-1}|a_{h-2}t} = 0.$$

Because of (4.16) this becomes

$$E_{p_h|a_{h-2}wt} - E_{p_h|a_{h-2}tw} = 0,$$

which vanishes identically by the very definition of the E's. Hence (4.11) is a consequence of (4.16). These last two relations are given by Mayer (M, p. 293) but his proof is from a somewhat different point of view.

COROLLARY. *If the matrix E_h is of rank one, the h-th curvature tensor vanishes identically.*

For since E_h is of rank one at any point $P \subset U$, we may write

$$E_{a_h|b_h} = \lambda_{a_h}\lambda_{b_h}.$$

Then (5.3) may be written

$$B^{s_h}_{p_h tw}\lambda_{s_h}\lambda_{a_h} + B^{s_h}_{a_h tw}\lambda_{s_h}\lambda_{p_h} = 0. \tag{5.11}$$

If we put $a_h = p_h$, the result is

$$B^{s_h}_{p_h tw}\lambda_{s_h}\lambda_{p_h} = 0.$$

But at least one λ_{p_h} (say λ_{q_h}) is not zero. Then the corresponding $B^{s_h}_{q_h tw}\lambda_{s_h} = 0$, and consequently for any a_h, $B^{s_h}_{q_h tw}E_{s_h|a_h} = 0$. But from this result and (5.11) it follows that

$$B^{s_h}_{a_h tw}\lambda_{s_h}\lambda_{q_h} = 0, \qquad B^{s_h}_{a_h tw}\lambda_{s_h} = 0. \tag{5.12}$$

Consequently all $B_{p_h|a_h|tw} = 0$ in U. This result has been obtained in a different manner by Burstin.[5]

6. **Determination of the Γ's in terms of the *E*'s.** We now turn our attention to equations (4.11)-(4.17) and deduce from them a set of necessary conditions on the coefficients of the fundamental forms. We know that in each $V(P)$ there exists a solution of (4.16), i.e.,

$$\Gamma^{d_h}_{b_{h+1}t}E_{d_h|a_h} = -E_{a_h t|b_{h+1}}, \tag{6.1}$$

where these are considered as linear equations in the Γ's. Hence we have the

MATRIX CONDITION I. *The augmented matrices $\bar{E}_h \equiv \| E_{d_h|a_h}; E_{a_h t|b_{h+1}} \|$ are of rank r_h in U for every choice of t and b_{h+1}.*

Similar linear equations in $\Gamma^{a_h}_{b_h t}$ may also be derived. Before proceeding to the general case, we illustrate the method by taking $h = 1$. From (4.17) it follows that

$$2\Gamma^{\gamma}_{\alpha\tau}E_{\gamma|\beta} = \frac{\partial E_{\alpha|\beta}}{\partial x^{\tau}} + \frac{\partial E_{\beta|\tau}}{\partial x^{\alpha}} - \frac{\partial E_{\tau|\alpha}}{\partial x^{\beta}}. \tag{6.2}$$

Since the determinant $|E_{\gamma|\beta}| \neq 0$, these may always be solved for $\Gamma^{\gamma}_{\alpha\tau}$. Thus the $\Gamma^{\gamma}_{\alpha\tau}$ turn out to be the Christoffel symbols of the second kind.

[5] C. Burstin, *Beiträge zur mehrdimensionalen Differentialgeometrie*, Monatshefte für Math. und Physik, vol. 36 (1929), p. 114.

For $h = 2$ we have from (4.17) and (4.15):

$$
\begin{aligned}
2E_{\sigma\rho|\gamma\delta}\Gamma^{\sigma\rho}_{\alpha\beta\epsilon} &= 2E_{\sigma\rho|\gamma\delta}(\Gamma^{\sigma}_{\alpha\epsilon}\delta^{\rho}_{\beta} + \Gamma^{\sigma}_{\beta\epsilon}\delta^{\rho}_{\alpha}) + E_{\alpha\beta|\gamma\delta,\epsilon} + E_{\beta\gamma|\delta\epsilon,\alpha} \\
&\quad + E_{\gamma\delta|\epsilon\alpha,\beta} - E_{\delta\epsilon|\alpha\beta,\gamma} - E_{\epsilon\alpha|\beta\gamma,\delta} \equiv \varphi_{\gamma\delta|\alpha\beta\epsilon},
\end{aligned}
\tag{6.3}
$$

where $\varphi_{\gamma\delta|\alpha\beta\epsilon}$ are thus defined. We know that these equations, considered as linear in the Γ's, have a solution in every $V(P)$. Thus we have the following

MATRIX CONDITION II. *For every choice of* $(\alpha\beta\epsilon)$ *the augmented matrices* $E'_2 \equiv \| E_{\sigma\rho|\gamma\delta}; \varphi_{\gamma\delta|\alpha\beta\epsilon} \|$ *are of rank* r_2 *in every* $V(P)$.

It will be noticed that these conditions are closely related to the Codazzi equations of the classical theory.

In order to extend these results to a general h, we define the operator Q to indicate the sum of terms obtained by cyclic interchange of $(\alpha_1 \cdots \alpha_h \beta_1 \cdots \beta_h \tau)$, a plus sign being attached to a term if the last index in the preceding bracket is τ or any α, and a negative sign if it is any β. Thus the general case of (6.2) and (6.3) is

$$
\begin{aligned}
2E_{\gamma_1\cdots\gamma_{h-1}\delta|\beta_1\cdots\beta_{h-1}\lambda}\Gamma^{\gamma_1\cdots\gamma_{h-1}\delta}_{\alpha_{h-1}\epsilon\tau} &= 2(\Gamma^{\gamma_1\cdots\gamma_{h-1}}_{\alpha_1\cdots\alpha_{h-1}\tau}\delta^{\delta}_{\epsilon} \\
&\quad + \Gamma^{\delta}_{\epsilon\tau}\delta^{\gamma_1\cdots\gamma_{h-1}}_{\alpha_1\cdots\alpha_{h-1}})\, E_{\gamma_1\cdots\gamma_{h-1}\delta|\beta_1\cdots\beta_{h-1}\lambda} + Q\Big(\frac{\partial E_{\alpha_1\cdots\alpha_{h-1}\epsilon|\beta_1\cdots\beta_{h-1}\lambda}}{\partial x^{\tau}} \\
&\quad - E_{\alpha_1\cdots\alpha_{h-1}\epsilon|\mu_1\cdots\mu_{h-1}\lambda}\Gamma^{\mu_1\cdots\mu_{h-1}}_{\beta_1\cdots\beta_{h-1}\tau} - E_{\mu_1\cdots\mu_{h-1}\epsilon|\beta_1\cdots\beta_{h-1}\lambda}\Gamma^{\mu_1\cdots\mu_{h-1}}_{\alpha_1\cdots\alpha_{h-1}\tau} \\
&\quad - E_{\alpha_1\cdots\alpha_{h-1}\rho|\beta_1\cdots\beta_{h-1}\lambda}\Gamma^{\rho}_{\epsilon\tau} - E_{\alpha_1\cdots\alpha_{h-1}\epsilon|\beta_1\cdots\beta_{h-1}\rho}\Gamma^{\rho}_{\lambda\tau}\Big).
\end{aligned}
\tag{6.4}
$$

In our briefer notation this may be written

$$
2E_{c_{h-1}d|b_{h-1}q}\Gamma^{c_{h-1}d}_{a_{h-1}pt} = 2(\Gamma^{c_{h-1}}_{a_{h-1}t}\delta^{d}_{p} + \Gamma^{d}_{pt}\delta^{c_{h-1}}_{a_{h-1}})E_{c_{h-1}d|b_{h-1}q} + Q(E_{a_{h-1}p|b_{h-1}q;t}),
\tag{6.5}
$$

or

$$
2E_{c_h|b_h}\Gamma^{c_h}_{a_h t} = \varphi_{b_h|a_h t},
$$

where $\varphi_{b_h|a_h t}$ are so defined and we have written c_h for $c_{h-1}d$, etc. Since we know that the Γ's exist which satisfy these equations, we get

MATRIX CONDITION II (*General case*). *The augmented matrices* $E'_h \equiv \| E_{c_h|b_h}; \varphi_{b_h|a_h t} \|$ *are of rank* r_h *in every* $V(P)$ *for each choice of* $(a_h t)$.

As we have stated them, the matrix conditions I involve $E_{a_h|b_h}$ and $E_{a_{h+1}|b_{h+1}}$ only. On the contrary, conditions II involve their first derivatives, and $\Gamma^{c_{h-1}}_{a_{h-1}t}$, and thus are not expressed as conditions on the fundamental forms. However, $\Gamma^{c_{h-1}}_{a_{h-1}t}$ is given by (6.5) in terms of $E_{a_{h-1}|b_{h-1}}$ and $\Gamma^{c_{h-2}}_{a_{h-2}t}$. By repeating this process we finally *can eliminate all* Γ'*s and may regard the matrix conditions* II *as involving only the fundamental forms.*

In a similar manner it follows that the Gauss equations (4.13) and (4.14) can be expressed in terms of the coefficients of the fundamental forms and their derivatives only. It is conceivable that in eliminating the Γ's by means of (6.5) the arbitrary θ's might enter into the resulting expressions. Direct examination of the formulas involved, however, shows that this is not the case. We have thus proved

THEOREM II. *The matrix conditions* I *and* II *and the equations* (4.13) *and* (4.14) *are conditions on the coefficients of the fundamental forms which are necessarily satisfied if the forms define an imbedding of a Riemann space.*

We now turn to the converse problem and determine conditions which are sufficient for these forms to define such an imbedding. We assume at the outset that the $E_{a_h|b_h}$ satisfy (1) of Theorem I and that they also satisfy the conditions stated in Theorem II. We define a set of Γ's in each $V(P)$ by the following recursive process. For $h = 1$, (6.2) define $\Gamma^{\alpha}_{\beta\gamma}$ uniquely as functions of class C^{m-1}. These actually satisfy all equations in (4.11)-(4.17) for $h = 1$, for the only equations which occur in this case are (4.15) and (4.17), and these are indeed satisfied identically because of (6.2).

For $h = 2$, we define $\Gamma^{\sigma\rho}_{\alpha\beta\epsilon}$ by means of (6.3). This is possible since the right hand sides contain only the given $E_{\alpha\beta|\gamma\delta}$, their derivatives, and a previously determined set of $\Gamma^{\alpha}_{\beta\gamma}$. Since the matrix E_2 is of rank r_2, a solution of (6.3) exists and is determined to within an additive function $\theta^{\sigma\rho}_{\alpha\beta\epsilon}$ which is a solution of (4.5). These solutions are defined as functions of x in each $V(P)$, and the additive functions can be so chosen that they are of class C^{m-2}. In a similar fashion $\Gamma^{\alpha}_{\beta\gamma\delta}$ are defined as functions of class C^{m-1} in U by equations (6.1). It remains to investigate whether the Γ's so defined actually satisfy (4.11)-(4.17) when $h = 2$. If we substitute for $\Gamma^{\sigma\rho}_{\alpha\beta\epsilon}$ in (4.15) and then use (4.13), we obtain

$$B_{\delta\alpha\beta\epsilon,\gamma} + B_{\gamma\delta\beta\epsilon,\alpha} + B_{\alpha\gamma\beta\epsilon,\delta} = 0.$$

But this is Bianchi's identity, and so our Γ's satisfy (4.15). Similarly the result of substituting for $\Gamma^{\alpha\epsilon}_{\beta\gamma\delta}$ in (4.17) vanishes because of (4.13) and the identity (5.9). Since (4.16) are satisfied by the very definition of $\Gamma^{\alpha}_{\beta\gamma\delta}$, and since (4.11) and (4.12) are consequences of the other equations, it follows that if $E_{\alpha|\beta}$ and $E_{\alpha\beta|\gamma\delta}$ satisfy the Gauss equations and if E_2' is of rank r_2, it is possible to find $\Gamma^{\alpha}_{\beta\gamma\delta}$ and $\Gamma^{\sigma\rho}_{\alpha\beta\epsilon}$ which satisfy (4.11)-(4.17) for $h = 2$.

For a general value of h we solve (6.1) and (6.5) for $\Gamma^{a_h-1}_{b_h t}$ and $\Gamma^{a_h}_{b_h t}$ in terms of the coefficients of the fundamental forms and previously determined Γ's. And as for the case $h = 2$, it can be shown that these new Γ's satisfy (4.11)-(4.17) as a consequence of the identities of §5. Hence we have

THEOREM III. *Let there be given a set of tensors* $E_{\alpha_1\cdots\alpha_h|\beta_1\cdots\beta_h}$ $(h = 1, \cdots, m)$ *which are defined as functions of class* C^{m-h+1} *in* U, *where* E_1 *is positive definite in* U, *and* E_h $(h = 2, \cdots, m)$ *is positive semi-definite of constant rank* r_h *in* U *and* $r_2 + \cdots + r_m = p$. *Then quantities* $\Gamma^{\alpha_1\cdots\alpha_h}_{\beta_1\cdots\beta_h\tau}$ *and* $\Gamma^{\alpha_1\cdots\alpha_{h-1}}_{\beta_1\cdots\beta_h\tau}$ *can be found which will satisfy the equations* (4.11)-(4.17) *in every* $V(P) \subset U$ *if and only if the Gauss equations are satisfied in* U, *the m-th curvature tensor vanishes in* U, *and the matrices* $\bar{E}_h$ *and* E_h' *are each of rank* r_h *in* U. *The* $\Gamma^{\alpha_1\cdots\alpha_h}_{\beta_1\cdots\beta_h\tau}$ *so defined are of class* C^{m-h} *and the* $\Gamma^{\alpha_1\cdots\alpha_{h-1}}_{\beta_1\cdots\beta_h\tau}$ *are of class* C^{m-h+1} *respectively.*

7. **Solution of the Frenet equations in a restricted neighborhood.** We now suppose that a set of E's are given which satisfy the conditions of the preceding theorem. In a definite $V(P)$ we define the Γ's by the method just described.

Consider further a spherical neighborhood $W(P)$ covered by a single coördinate system and such that $V(P) \supset W(P) \supset P$, i.e., one such that its map in the arithmetic space is the interior of a sphere. We now seek to integrate the mixed system composed of the Frenet equations (4.1) and the algebraic equations (3.2) and (3.3) in the neighborhood $W(P)$.

For convenience let us denote (4.1) by

$$\frac{\partial}{\partial x^\alpha}(y^i) = Y^i_\alpha; \cdots; \frac{\partial Y^i_{a_h}}{\partial x^\gamma} = \varphi^k_{a_h\gamma} Y^i_k \qquad \begin{pmatrix} i = 1, \cdots, n+p; \\ h = 1, \cdots, m; \\ k = a_{h-1}, a_h, a_{h+1} \end{pmatrix}. \tag{7.1}$$

Let the coördinates of P be x_0, and assign the values $(y^i)_0$; $(Y^i_{a_h})_0$ to the unknowns at P such that (3.2) and (3.3) are satisfied when $x = x_0$. It is clear that this choice of initial values is always possible, for the matrix E is by hypothesis a symmetric, positive semi-definite matrix of rank $n + p$. Therefore a canonical representation of the form (3.2) and (3.3) is possible at P.[6]

Moreover, any two possible sets of initial values $(Y^i_{a_h})_0$ are connected by the relation

$$(Y^i_{a_h})_0 = a^i_j (Y'^j_{a_h})_0, \tag{7.2}$$

where the a's are an orthogonal matrix of constants. The values of $(y^i)_0$ are completely arbitrary.

Picking out one such set of initial values we now seek to find the corresponding values of the unknowns at another point Q of coördinates $x_1 \subset W(P)$. To do this we join P and Q with an arbitrary curve C, lying in $W(P)$, defined by $x^\alpha = f^\alpha(t)$, where $f^\alpha(0) = x^\alpha_0$; $f^\alpha(1) = x^\alpha_1$ and $f^\alpha(t)$ are single valued and of class C^1 for $0 \leqq t \leqq 1$. There always exists such a curve, for in particular we may take $f^\alpha(t) = x^\alpha_0 + (x^\alpha_1 - x^\alpha_0)t$. Now integrate

$$\frac{dy^i}{dt} = Y^i_\alpha \frac{dx^\alpha}{dt}; \cdots; \frac{dY^i_{a_h}}{dt} = \varphi^k_{a_h\gamma}[x(t)] Y^i_k \frac{dx^\gamma}{dt} \tag{7.3}$$

along C for the above set of initial values. The quantities $y^i(t)$, $Y^i_{a_h}(t)$ thus defined satisfy (3.2) and (3.3) for $t = 0$ by hypothesis, and indeed they satisfy (3.2) and (3.3) all along C. For let

$$[\alpha_1 \cdots \alpha_h \quad \beta_1 \cdots \beta_h] \equiv Y^i_{\alpha_1 \cdots \alpha_h} Y^i_{\beta_1 \cdots \beta_h} - E_{\alpha_1 \cdots \alpha_h | \beta_1 \cdots \beta_h}; \tag{7.4}$$

$$[\alpha_1 \cdots \alpha_h \quad \beta_1 \cdots \beta_k] \equiv Y^i_{\alpha_1 \cdots \alpha_h} Y^i_{\beta_1 \cdots \beta_k} \qquad (h \neq k). \tag{7.5}$$

Then (3.2) and (3.3) may be written

$$[\alpha_1 \cdots \alpha_h \quad \beta_1 \cdots \beta_l] = 0 \qquad (h, l = 1, \cdots, m). \tag{7.6}$$

Differentiating (7.4) and (7.5) along C and then applying (7.4) and (7.5) to the result, we find that the derivatives of the bracket expressions are linear combinations of the bracket expressions themselves, the coefficients involving

[6] Cf. Duschek-Mayer, *Lehrbuch der Differentialgeometrie*, II, p. 25.

the Γ's and dx^α/dt, plus terms which vanish because of (4.16) and (4.17). But this shows at once that (7.6) is satisfied all along C.

THEOREM IV. *The solutions* $Y^i_{a_h}(t)$ *of* (7.3) *whose initial values satisfy* (3.2) *and* (3.3) *for* $x = x_0$ *satisfy* (3.2) *and* (3.3) *for all points of* C.

We have assumed that the conditions (4.11)-(4.17) are satisfied in $W(P)$. Moreover, we know from §3 that if (4.11)-(4.15) hold, then the integrability conditions (4.6)-(4.10) are satisfied provided that (3.2) and (3.3) are true. Hence we have

THEOREM V. *The integrability conditions* (4.6)-(4.10) *of the Frenet equations are satisfied along the curve* C *by the solutions of* (7.3), *whose initial values satisfy* (3.2) *and* (3.3) *for* $x = x_0$.

In order to prove that the values of the y^i and $Y^i_{a_h}$ thus defined at Q are genuine functions of the x's, it is necessary to show that these values are independent of the curve C along which we have integrated. Join P and Q with some other curve C_1 of class C^1 lying in $W(P)$ whose equation is $x^\alpha = f_1^\alpha(t)$, where f_1^α have the same properties as f^α.[7] Since $W(P)$ is a spherical neighborhood, C and C_1 can be imbedded in a one-parameter family of curves $x^\alpha = G^\alpha(t, p)$, joining P and Q and lying in $W(P)$. The $G^\alpha(t, p)$ are assumed to possess the following continuous derivatives

$$\frac{\partial G}{\partial t}; \qquad \frac{\partial G}{\partial p}; \qquad \frac{\partial^2 G}{\partial t\, \partial p} = \frac{\partial^2 G}{\partial p\, \partial t}$$

for $0 \leqq t \leqq 1$ and $0 \leqq p \leqq 1$ and to be such that $G^\alpha(t, 0) = f^\alpha(t)$ and $G^\alpha(t, 1) = f_1^\alpha(t)$; $G^\alpha(0, p) = x_0$; and $G^\alpha(1, p) = x_1^\alpha$. In particular, we may put $G^\alpha(t, p) = f^\alpha(t) + p[f_1^\alpha(t) - f^\alpha(t)]$.

Consider the equations

$$(7.7) \qquad \frac{dy^i}{dt} = Y^i_\alpha[G(p, t)]\,\frac{\partial G^\alpha}{\partial t}; \cdots ; \frac{dY^i_{a_h}}{dt} = \varphi^k_{a_h\gamma}[G(p, t)]\, Y^i_k\, \frac{\partial G^\gamma}{\partial t}$$

along any curve of the family, i.e., for a particular value of p. If we take for the values of the unknown at P (corresponding to $t = 0$) the quantities which served as initial values for our integration of (7.3), the equations (7.7) define a set of functions $y^i(t, p)$, $Y^i_{a_h}(t, p)$ for all values of t and p in the specified ranges. Because of Theorem IV they also satisfy

$$(7.8) \qquad E_{a_h|b_h}[G(t, p)] = Y^i_{a_h}[G(t, p)]Y^i_{b_h}[G(t, p)].$$

For $t = 1$ we have the values of $y^i(1, p)$ and $Y^i_{a_h}(1, p)$ at Q as defined by the curve of the family of parameter p. We now show that these values are actually independent of p.

[7] The following treatment is based on that given by T. Y. Thomas, *Systems of total differential equations defined over simply connected domains*, Annals of Math., vol. 35 (1934), pp. 730–734. Thomas' treatment, however, applies only to completely integrable systems which do not involve additional algebraic equations.

Consider the equations

$$\frac{\partial y^i}{\partial t} = Y^i_\alpha(G)\,\frac{\partial G^\alpha}{\partial t}; \qquad \frac{\partial Y^i_{\alpha_h}}{\partial t} = \varphi^k_{\alpha_h\gamma}(G)\,Y^i_k(G)\,\frac{\partial G^\gamma}{\partial t}, \tag{7.9}$$

$$\begin{cases} \dfrac{\partial y^i}{\partial p} = Y^i_\alpha(G)\,\dfrac{\partial G^\alpha}{\partial p} + \sigma^i_0, \\[2mm] \dfrac{\partial Y^i_{\alpha_h}}{\partial p} = \varphi^k_{\alpha_h\gamma}(G)\,Y^i_k(G)\,\dfrac{\partial G^\gamma(t,p)}{\partial p} + \sigma^i_{\alpha_h}, \end{cases} \tag{7.10}$$

where σ^i_0 and $\sigma^i_{\alpha_h}$ are defined by the last equations. If we differentiate (7.9) with respect to p, and (7.10) with respect to t, and equate the right hand sides of the resulting equations, we obtain

$$\begin{aligned} \frac{\partial \sigma^i_0}{\partial t} &= \sigma^i_\gamma\,\frac{\partial G^\gamma}{\partial t} + \Phi^k_{0\gamma\delta}\,Y^i_k\,\frac{\partial G^\gamma}{\partial t}\,\frac{\partial G^\delta}{\partial p}; \\ \frac{\partial \sigma^i_{\alpha_h}}{\partial t} &= \sigma^i_k\varphi^k_{\alpha_h\gamma}\,\frac{\partial G^\gamma}{\partial t} + \Phi^k_{\alpha_h\gamma\delta}\,Y^i_k\,\frac{\partial G^\gamma}{\partial t}\,\frac{\partial G^\delta}{\partial p}, \end{aligned} \tag{7.11}$$

where $\Phi^k_{0\gamma\delta}$ and $\Phi^k_{\alpha_h\gamma\delta}$ are the bracketed expressions on the left sides of (4.6)-(4.10). From Theorem V we see that

$$\Phi^k_{0\gamma\delta}\,Y^i_k = 0, \qquad \Phi^k_{\alpha_h\gamma\delta}\,Y^i_k = 0$$

for all allowable t and p.

Now for $t = 0$, we have y^i, $Y^i_{\alpha_h}$, and $G^\alpha(0, p)$ by hypothesis independent of p; hence from (7.10) it follows that $\sigma^i_0 = 0$ and $\sigma^i_{\alpha_h} = 0$ for $t = 0$. And so from (7.11) we have that $\sigma^i_0 = 0$ and $\sigma^i_{\alpha_h} = 0$ for all allowable values of t and p, and in particular for $t = 1$. We have also by hypothesis that $G(1, p)$ are independent of p, and hence from (7.10) it follows that the solutions of (7.7) at $t = 1$ are independent of p.

We therefore define $y^i(x)$, $Y^i_{\alpha_h}(x)$ as functions of x in $W(P)$ whose values at any point $Q \subset W(P)$ are obtained from a given set of initial values satisfying (3.2) *and* (3.3) *at P by integrating along any curve of class C^1 joining P and Q and lying in $W(P)$. These functions will satisfy* (3.2) *and* (3.3) *in $W(P)$.*

The proof that these functions are unique and that they are actually solutions of the differential equations is given by Thomas.[7] It is also shown there that if $Q \subset W(P)$ and if initial values are chosen satisfying (3.2) and (3.3) at Q, there exists a unique solution of the mixed system in $W(P)$ having the given initial values at Q. To find this solution, it is only necessary to integrate along a curve from Q to P, thus determining values of the functions at P. With these as initial values the system can be integrated in $W(P)$, and it is shown that the resulting solution has the required initial values at Q. We shall need this property in the next section.

THEOREM VI. *Let* (1) *there be given a set of functions $E_{\alpha_1\cdots\alpha_h|\beta_1\cdots\beta_h}$ $(h = 1, \cdots, m)$ of class C^{m-h+1} in a spherical neighborhood $W(P)$ such that E_1 is positive definite in $W(P)$; E_h is positive semi-definite of constant rank r_h $(h = 2, \cdots, m)$*

in $W(P)$ *and* $r_2 + \cdots + r_m = p$; (2) *the matrices* E'_h *and* $\bar{E}_h$ *be of rank* r_h *in* $W(P)$; (3) *the* $E_{\alpha_1 \cdots \alpha_h | \beta_1 \cdots \beta_h}$ *and their derivatives satisfy the Gauss equations in* $W(P)$; (4) *the m-th curvature tensor vanish in* $W(P)$.

Then and only then do the Frenet equations (4.1) *and their associated algebraic equations* (3.2) *and* (3.3) *have a unique solution* $y^i(x)$, $Y^i_{\alpha_1 \cdots \alpha_h}(x)$ *for all* $x \subset W(P)$ *for each set of initial values* $(y^i)_0$ *and* $(Y^i_{\alpha_1 \cdots \alpha_h})_0$ *which satisfy* (3.2) *and* (3.3) *at any point* $Q \subset W(P)$. *The Riemann space defined by* $E_{\alpha | \beta}$ *in* $W(P)$ *is thus imbedded in a Euclidean* E_{n+p}.

8. **Solution of the Frenet equations in an open simply connected domain.** Next we consider an open simply connected domain U of a coördinate manifold. Suppose that functions $E_{a_h | b_h}$ $(h = 1, \cdots, m)$ satisfying the hypothesis of the last theorem are defined throughout U. Then each point $P \subset U$ is contained in at least one spherical neighborhood $W(P) \subset U$ within which the Γ's may be defined as in §6. If two such neighborhoods $W(P)$ and $W(P')$ have points in common, from §4 we see that the $(\Gamma)_P$ defined in $W(P)$ and the $(\Gamma)_{P'}$ defined in $W(P')$ satisfy the equations

$$
\begin{aligned}
&E_{a_h | b_h}[(\Gamma^{b_h}_{a_{h+1}})_P - (\Gamma^{b_h}_{a_{h+1}})_{P'}] = 0, \\
&E_{a_{h-1} | b_{h-1}}[(\Gamma^{b_{h-1}}_{a_{h+1}})_P - (\Gamma^{b_{h-1}}_{a_{h+1}})_{P'}] = 0
\end{aligned}
\tag{8.1}
$$

in $W(P) \cap W(P')$. We shall now show that the conclusions of the theorem of §7 can be extended to the entire neighborhood U.

Pick out any point $P \subset U$ of coördinates x_0 and assign initial values $(y^i)_0$, $(Y^i_{a_h})_0$ to the unknowns at P which satisfy (3.2) and (3.3) at P. In order to determine the corresponding value at another point $Q \subset U$, join P and Q with a simple curve C of class C^1. This is possible, since U is connected. Because of the Heine-Borel Theorem, C can be covered by a finite set of the spherical neighborhoods $W_1, \cdots, W_f$, where $W_1 = W(P)$ and $W_f = W(Q)$. Let $V_1, \cdots, V_f$ be points on C such that $V_g \subset W_g \cap W_{g+1}$ $(g = 1, \cdots, f-1)$ and $V_f = Q$. Integrating along C from P to V_1, we obtain values $(y^i)_1$ and $(Y^i_{a_h})_1$ at V_1 corresponding to the given initial values. Then taking these values at V_1 as initial values integrate in W_2 from V_1 to V_2 and continue this process until finally by integration from V_{f-1} to Q values $(y^i)_Q$ and $(Y^i_{a_h})_Q$ are defined at Q.

Suppose next that P and Q are joined by another simple curve of class C^1, say C_1, which passes through the same set of neighborhoods $W_1, \cdots, W_f$ in their natural order. Select points $X_1, \cdots, X_f$ on C_1 such that $X_g \subset W_g \cap W_{g+1}$ $(g = 1, \cdots, f-1)$ and $X_f = Q$. Now by the above process define values $(y^i)'_Q$ and $(Y^i_{a_h})_Q$ at Q corresponding to the curve C_1 for the given initial values at P.

Let V_g and X_g $(g = 1, \cdots, f-1)$ be joined by some curve of class C_1, say S_g, lying in $W_g \cap W_{g+1}$. Integrating (7.3) along S_g from V_g to X_g for the values of the Γ's, $(\Gamma)_g$, defined in W_g, and for initial values satisfying (3.2)

and (3.3), we obtain a set of $y^i(t)$ and $Y^i_{a_h}(t)$ for each point of S_g. Now (7.3) are actually equations of the form

$$(8.2)\qquad \begin{aligned} \frac{dy^i}{dt} &= Y^i_\alpha \frac{dx^\alpha}{dt}, \\ \frac{d}{dt}(Y^i_{a_h}) &= [(\Gamma^{b_{h-1}}_{a_h w})_g Y^i_{b_{h-1}} + (\Gamma^{b_h}_{a_h w})_g Y^i_{b_h} + Y^i_{a_h w}] \frac{dx^w}{dt}. \end{aligned}$$

Now substitute the $y^i(t)$ and $Y^i_{a_h}(t)$ obtained by the above integration along S_g into the equations (8.2) in which the $(\Gamma)_g$ are replaced by the $(\Gamma)_{g+1}$ defined in W_{g+1}, and subtract the resulting equations from (8.2). We obtain

$$(8.3)\quad [(\Gamma^{b_{h-1}}_{a_h w})_g - (\Gamma^{b_{h-1}}_{a_h w})_{g+1}] Y^i_{b_{h-1}} \frac{dx^w}{dt} + [(\Gamma^{b_h}_{a_h w})_g - (\Gamma^{b_h}_{a_h w})_{g+1}] Y^i_{b_h} \frac{dx^w}{dt} = 0.$$

Because of (8.1) and the fact that (3.2) and (3.3) are satisfied along S_g, we see that (8.3) are satisfied identically in t. *Hence integration along any S_g gives the same result, whether S_g is regarded as a curve of W_g or as a curve of W_{g+1}.*

Next we observe that the values of the unknowns at X_1 are the same whether the integration has proceeded from P to X_1 along C_1 or from P to V_1 along C and then along S_g from V_1 to X_1; for all these curves lie in $W(P)$. We indicate this symbolically by

$$(8.4)\qquad PX_1 \approx PV_1X_1.$$

Now suppose, in order to establish an induction, that the values at X_{g-1} are the same whether we have proceeded along C_1 to X_{g-1} or along C to V_{g-1} and then along S_{g-1} from V_{g-1} to X_{g-1}. That is, we assume that

$$(8.5)\qquad PX_{g-1} \approx PV_{g-1}X_{g-1}.$$

Then the values at X_g as found by proceeding along C_1 from P to X_g are the same as those obtained from considering the curve C from P to V_g and then the curve S_g from V_g to X_g. That is,

$$(8.6)\qquad PX_g \approx PV_gX_g.$$

For to values at X_{g-1} given by integration along C_1 from P to X_{g-1}, the corresponding values at X_g are the same for the integration along $X_{g-1}X_g$ as for the integration along $X_{g-1}V_{g-1}V_gX_g$, since all curves involved lie in W_g; i.e.,

$$(8.7)\qquad X_{g-1}X_g \approx X_{g-1}V_{g-1}V_gX_g.$$

Combining (8.7) and (8.5), we establish (8.6). Since the hypothesis of the induction holds for $g = 1$ because of (8.4), the relation (8.6) holds for any g, and in particular for $g = f$, i.e., at Q. Therefore *the values $(y^i)_Q$ and $(Y^i_{a_h})_Q$ determined by the curve C covered by $W_1, \cdots, W_f$ are the same as those determined by any other simple curve of class C^1 lying in $W_1, \cdots, W_f$ and passing through them in their natural order.*

Finally, consider any simple curve C_2 of class C^1, lying in U, joining P and Q, and distinct from C. As above C_2 may be covered with a finite number of neighborhoods and gives rise to a set of values $\overline{(y^i)}_Q$ and $\overline{(Y^i_{\alpha_k})}_Q$ corresponding to the given set of values at P. Suppose that C and C_2 are imbedded in a one-parameter family of curves joining P to Q and such that for the parameter value $p = 0$ we have the curve C and for $p = 1$ we have C_2. Assume also that the coördinates x^α on curves of the family are continuous functions of p for $0 \leqq p \leqq 1$. Hence if some member of the family, say $C_{\bar{p}}$, of parameter $\bar{p}$ is covered by a finite set of neighborhoods, $Y_1, \cdots, Y_f$, there exists a $\delta > 0$ such that for $|p - \bar{p}| < \delta$ all curves of parameter p lie in $Y_1, \cdots, Y_f$ and pass through them in their natural order. Because of this property and the last italicized statement, the $[y^i(p)]_Q$ and $[Y^i_{\alpha_k}(p)]_Q$ are differentiable with respect to p for $0 \leqq p \leqq 1$, and in fact the derivatives are zero for each value of p. Hence the unknowns are uniquely determined at Q for given initial values satisfying (3.2) and (3.3) at P by integrating along any simple curve of class C^1 joining P to Q and lying in U.

THEOREM VII. *Theorem VI remains true if the spherical neighborhood is replaced by an open simply connected domain.*

From (7.2) we have the fact that any two sets of initial values of Y^i_α which satisfy (3.2) and (3.3) must have the relation

$$(Y^i_\alpha)_0 = a^i_j\,(Y^j_\alpha)'_0, \tag{8.8}$$

where $\|a^i_j\|$ is orthogonal. Hence the corresponding y's satisfy

$$y^i = a^i_j y'^j + b^i, \tag{8.9}$$

where b^i are constants determined by the given initial values of the y^i and $y^{i\prime}$. *Therefore, the above imbedding is determined to within a motion in the Euclidean space.*

PRINCETON UNIVERSITY.

REMARK ON THE THEOREM OF GREEN

By S. Bochner

The theorem of Green is an identity between an integral over a compact region R of an orientable locally Euclidean separable space S of class two[1] whose dimension will be denoted by n, and an integral over the boundary B of R, in case B is formed by "hypersurfaces":

$$(1) \qquad \int_R \operatorname{div} \lambda dv = \int_B \lambda^i \Phi_i \, d\omega.$$

If R is closed, that is, compact and equal to S, thus having "no boundary", the integral over the boundary is to be put equal to zero:

$$(2) \qquad \int_S \operatorname{div} \lambda dv = 0.$$

If R is contained in one coördinate neighborhood, the proof of (1) is comparatively simple, provided the boundary B is sufficiently smooth with respect to the coördinate system.[2] But the passage from the local case to a domain R in the large is rather laborious. It requires a cellular subdivision of R into sufficiently small subregions whose boundary is sufficiently smooth, an application of the local theorem to each subregion, and finally a justification of the mutual cancellation of the boundary terms arising from the artificial cellular partitions. Now this procedure is much too complicated and heavy in the case of formula (2) or in case of formula

$$(3) \qquad \int_R \operatorname{div} \lambda dv = 0 \qquad\qquad (\lambda^i \equiv 0 \text{ on } B).$$

We want to show that these two formulas can be deduced in a much simpler fashion, even avoiding any complication that might be inherent to the local theorem itself. In particular we shall eliminate from the proof the concept of $[(n-1)$-dimensional$]$ volume on the boundary B.

Our space S being of class two, we can consider on it tensors of class one and tensor densities of class one. In particular, we assume the existence of a positive (non-vanishing) scalar density which, as in the special case of Riemann spaces, will be denoted by $\sqrt{g}$; in going over from coördinates $(x_1, \cdots, x_n)$ to coördinates $(y_1, \cdots, y_n)$, the quantity $\sqrt{g}$ is to be multiplied by the jacobian

Received December 19, 1936.

[1] That is, a space allowing local coördinate transformations with continuous first and second partial derivatives and a positive (non-vanishing) Jacobian.

[2] A. Duschek and W. Mayer, *Differentialgeometrie*, 1930, vol. II, p. 237

$|\partial x/\partial y|$.[3] With such a scalar density we can form an invariant volume element

$$dv \equiv dv_x = \sqrt{g}\, dx_1 \cdots dx_n. \tag{4}$$

It gives rise to a regular Lebesgue measure and a Lebesgue integral as in the Euclidean case $\sqrt{g} = 1$; the only material difference being that in the case of every closed space S the total measure of S is finite.

If λ^i is a contravariant vector, the quantity

$$\operatorname{div} \lambda = \frac{1}{\sqrt{g}} \frac{\partial(\sqrt{g}\, \lambda^i)}{\partial x_i} \tag{5}$$

is an absolute scalar. This follows easily from the fact that the $(n - 1)$-dimensional minors A_{ij} (with appropriate signs) of the matrix

$$\left\| \frac{\partial x_i}{\partial y_j} \right\| \tag{6}$$

satisfy the relation

$$\frac{\partial A_{ij}}{\partial y_j} = 0 \qquad (i = 1, \cdots, n). \tag{7}$$

Hence we can form the integral of div λ over any bounded measurable set A. If A is contained in a coördinate neighborhood, then

$$\int_A \operatorname{div} \lambda dv = \int_A \frac{\partial(\sqrt{g}\, \lambda^i)}{\partial x_i} dx_1 \cdots dx_n. \tag{8}$$

Suppose now that A is a rectangle $a_\nu < x_\nu < b_\nu$, $\nu = 1, \cdots, n$, and that λ^i vanishes on the boundary of A. In this case, for each i,

$$\int_{a_i}^{b_i} \frac{\partial(\sqrt{g}\, \lambda^i)}{\partial x_i} dx_i = 0,$$

and therefore

$$\int_A \operatorname{div} \lambda dv = 0. \tag{9}$$

Also, the integral of div λ is 0 over any open set on which λ^i vanishes. Hence we obtain the first result that (2) is true if λ^i vanishes outside some "rectangle" A. If S is closed, and hence bicompact, we can cover it by a finite number of neighborhoods $U_1, \cdots, U_m$, whose closures are contained in "rectangles" $A_1, \cdots, A_m$ respectively. Corresponding to each μ, $\mu = 1, \cdots, m$, we can easily find a neighborhood V_μ between U_μ and A_μ and a non-negative scalar function φ_μ of class one in A_μ which is $\geqq 1$ in U_μ and 0 outside V_μ. Completing φ_μ by values 0 outside A_μ we have, throughout S, $\varphi_1 + \cdots + \varphi_m \geqq 1$. Hence the functions of class one

[3] O. Veblen, *Invariants of Quadratic Differential Forms*, 1927, pp. 19–25.

$$\psi_\mu = \frac{\varphi_\mu}{\varphi_1 + \cdots + \varphi_m} \tag{10}$$

have the following properties: each vanishes outside some rectangle and their sum is 1. For fixed μ we multiply the vector λ^i by the scalar ψ_μ. The resulting vector $\lambda^i_{(\mu)}$ vanishes outside a rectangle, and hence satisfies (2). But div $\lambda = \sum_{\mu=1}^{m} \operatorname{div} \lambda_{(\mu)}$, and this completes the proof of (2).

Formula (3) can be proved in a similar way provided λ^i is *strictly regular up to the boundary,* this expression meaning that, if completed by values 0 outside R, it is a vector of class one throughout S, so that (3) may be written

$$\int_S \operatorname{div} \lambda dv = 0 \qquad (\lambda^i \equiv 0 \text{ outside } R + B). \tag{11}$$

In fact, we can construct non-negative functions of class one $\varphi_1, \cdots, \varphi_m$, each vanishing outside some rectangle and whose sum is $\geqq 1$ for all points of $R + B$. The functions (10) will be of class one in some neighborhood U of $R + B$, although perhaps not throughout S. But λ^i vanishes outside $R + B$, and therefore the vectors $\lambda^i_{(\mu)}$ are again of class one, and we can again prove (3) from

$$\int_S \operatorname{div} \lambda dv = \sum_{\mu=1}^{m} \int_S \operatorname{div} \lambda_\mu \, dv. \tag{12}$$

We can make further statements if S is a Riemann space with a positive-definite fundamental tensor g_{ij} having the following property. *Corresponding to any point x^0 of S there exists a neighborhood $U(x^0)$ such that for any points x, y of $U(x^0)$ the function $\Omega(x, y)$, which is the square of the geodesic distance between x and y, has continuous first partial derivatives in $(x_1, \cdots, x_n; y_1, \cdots, y_n)$ and that, for some constant $C = C(x_0)$,*

$$\| \operatorname{grad}_x \Omega \|^2 \equiv g^{ij}(x) \frac{\partial \Omega}{\partial x_i} \frac{\partial \Omega}{\partial x_j} \leqq C\Omega(x, y). \tag{13}$$

The neighborhood $U(X^0)$ entering our condition can be assumed to be a sphere with center at x^0: $\Omega(x, x_0) < \rho$, $\rho > 0$. If x^0 transverses a set A whose closure is compact, then by the Heine-Borel theorem, this radius ρ and the constant $C(x^0)$ may be chosen uniformly with respect to A: $\rho = \rho(A)$, $C = C(A)$. If $0 < r < r_0 = \frac{1}{2}\rho(A)$, then A_r will denote the open set consisting of all points x for which $\Omega(x, x^0) < r$ for some $x^0 \subset A$. Let $\varphi(t)$ be any non-negative function of the real variable t which has a continuous first derivative in $0 \leqq t \leqq 3$, is positive for $0 \leqq t \leqq \frac{1}{2}$ and 0 for $1 \leqq t \leqq 3$. If $3r < r_0$, the expression

$$P_r(x) = \int_{A_r} \varphi\left(\frac{\Omega(x, y)}{r^2}\right) dv_y \Bigg/ \int_{A_{3r}} \varphi\left(\frac{\Omega(x, y)}{r^2}\right) dv_y \tag{14}$$

defines a non-negative function of class one in A_{3r} which has the value 1 on A and the value 0 outside A_{2r}. Completing $P_r(x)$ by the values 0 outside A_{3r}, we obtain a non-negative function of class one throughout S which is $\leqq 1$ everywhere, has the value 1 on A and the value 0 outside A_{2r}. Owing to the nature of our function $\varphi(t)$, the denominator is bounded from below by a constant multiple of r^n uniformly for all points x in A_r, and both the numerator and denominator are bounded from above by a constant multiple of r^n, for the given function $\varphi(t)$, and, likewise, if this function is replaced by $\varphi'(t)$. Forming now the partial derivatives of $P_r(x)$ and using (13), we find that

$$\text{(15)} \qquad \| \operatorname{grad}_x P_r(x) \| \leqq \frac{C_0}{r}, \text{ uniformly for } x \subset A_r.$$

We are now in a position to drop completely in formula (3) the requirement that λ^i be strictly regular up to the boundary and to interpret the boundary condition "$\lambda^i \equiv 0$ on B" as a very light condition requiring that $\lambda^i(x)$ have limit values 0 as x approaches the boundary B from within the region A.

THEOREM. *If R is an open region whose closure $R + B$ is compact and λ^i is a contravariant vector of class one in R, then*

$$\text{(16)} \qquad \int_R \operatorname{div} \lambda dv = 0$$

holds, provided

$$\text{(17)} \qquad \int_R | \operatorname{div} \lambda | \, dv < \infty$$

and

$$\text{(18)} \qquad \lim_{r \to 0} \frac{1}{r} \int_{B_r} \sqrt{g_{ij}\lambda^i\lambda^j} \, dv = 0.$$

Proof. Condition (17) simply states that our integral shall also be absolutely convergent and condition (18) requires that the "length" of $\lambda^i(x)$ shall "in average" tend to zero as x tends to the boundary, in the sense that the integral of $\| \lambda^i(x) \|$ over the set of points having a distance less than r from the boundary shall, as r tends to 0, be small in relation to the quantity r.

Now we shall give the proof itself. We consider the region R, the boundary B, and the exterior $\bar{R}$. In $B + \bar{R}$ we complete λ^i by values 0, and we multiply $\lambda^i(x)$ by the function

$$\text{(19)} \qquad Q_r(x) = 1 - P_r(x),$$

the function $P_r(x)$ being formed with respect to the compact set B. It is easily seen that $\lambda^i_{(r)} = Q_r \cdot \lambda^i$ is strictly regular up to the boundary and vanishes on B. Hence (16) holds for $\lambda^i_{(r)}$. All we have to prove now is the limit relation

$$\text{(20)} \qquad \int_R \operatorname{div} (\lambda - \lambda_{(r)}) dv \equiv \int_{R \cdot B_{2r}} \operatorname{div} (P_r \lambda) dv \to 0 \qquad \text{as } r \to 0.$$

But this follows from

$$\operatorname{div}(P_r\cdot\lambda) = P_r\cdot\operatorname{div}\lambda + \operatorname{grad} P_r\cdot\lambda$$

in connection with (15), (17), (18) and $|P_r(x)| \leqq 1$.

In the theorem we have just proved, the boundary may in part or totally consist of points which are limit points of R and not of $\bar{R}$; in other words, the theorem also takes care of the case in which the vector has singularities along lower dimensional manifolds. In particular it can be concluded that in Green's theorems (1), (2), (3) the vector λ^i may cease to be of class one along sufficiently smooth manifolds of dimension $\leqq n - 2$ as long as its length stays bounded in the neighborhood of these manifolds.

We can also make a statement regarding equation (1) itself. Let S be closed, and suppose that on some given boundary B it is possible to define an area element $d\omega$ and a vector Φ_i such that (1) holds for every vector λ^i which is defined and of class one in a neighborhood of $R + B$. Then the same is true for the exterior $\bar{R}$ of $R + B$, which means that

$$\int_{\bar{R}} \operatorname{div}\lambda\, dv = \int_B \lambda^i(-\Phi_i)\, d\omega \tag{21}$$

for every function λ^i which is defined and of class one in a neighborhood of $\bar{R} + B$, provided that the (n-dimensional) volume of the set B is 0. In fact, (21) follows immediately frcm (1) and (2) whenever λ^i is defined and of class one throughout S. But the values of the two integrals in (21) depend only on the values of λ^i on $\bar{R} + B$. Hence (21) will hold for a vector λ^i which, without altering its values on $\bar{R} + B$, can be modified into a vector which is of class one throughout S. Now if λ^i is of class one in a neighborhood of $\bar{R} + B$, this modification can be obtained by multiplying λ^i with a function $P_r(x)$ that belongs to the compact set $\bar{R} + B$ and by putting it 0 in other points.

PRINCETON UNIVERSITY.

ANALYTIC MAPPING OF COMPACT RIEMANN SPACES INTO EUCLIDEAN SPACE

By S. Bochner

Recently H. Whitney[1] has proved that an n-dimensional separable coördinate space S of class C_q $(1 \leqq q \leqq \infty)$ can be mapped topologically onto the Euclidean E_{2n+1} in such a manner that the mapping functions

$$t_\nu = t_\nu\,(x_1, \cdots, x_n) \qquad (\nu = 1, \cdots, 2n+1) \tag{1}$$

belong to class C_q on S and have a Jacobian of rank n throughout S; the quantities $x_1, \cdots, x_n$ in (1) are local coördinates on S varying with the neighborhood.

It is not known whether for an analytic space S the mapping functions (1) can be chosen analytic. It is the purpose of the present paper to point out that they can be so chosen provided S is compact and has an analytic Riemann metric.

The line of reasoning is very simple. With the fundamental tensor $g_{ij}(x)$, we form the Laplacian

$$\Delta\varphi = \frac{1}{g^{\frac{1}{2}}}\frac{\partial}{\partial x_i}\left(g^{\frac{1}{2}}\, g^{ij}\,\frac{\partial\varphi}{\partial x_j}\right) = g^{ij}\varphi_{,ij}\,. \tag{2}$$

In the Hilbert space H of all square integrable functions on S, the Laplacian is essentially the inverse of a completely continuous operator. Therefore the solutions φ of the equation

$$\Delta\varphi = \lambda\varphi \tag{3}$$

form a complete basis in H. But the solutions of (3) are analytic if the coefficients g_{ij} are so. Hence every function $t(x)$ on S is the limit in square mean of analytic functions. For differentiable functions on S we shall prove more; if $t(x)$ belongs to C_∞, then corresponding to any $\epsilon > 0$ there exists an analytic function $\psi^\epsilon(x)$ such that *the function* $t(x) - \psi^\epsilon(x)$ *and its gradient* differ by less than ϵ throughout S.[2] We now apply this approximation to the functions (1). It is easy to see that, for ϵ sufficiently small, the approximating transformation

$$t_\nu = \psi^\epsilon_\nu(x_1, \cdots, x_n) \tag{4}$$

will have the same mapping properties as the original transformation (1), which proves our assertion.

Received February 16, 1937.

[1] *Differentiable Manifolds*, Annals of Math., vol. 37 (1936), pp. 645–680; p. 654, Theorem 1.

[2] In order to prove this conclusion it would be sufficient to assume that $t(x)$ belongs to C_2. But the proof would become more elaborate.

The Laplacian $\Delta\varphi$ on a compact space S was treated by Hilbert (for $n = 2$)[3] and by G. Giraud.[4] In order to avoid too many references to these and other papers, which would have to be supplemented in minor points in any case, we shall reproduce a number of known details in the first half of this paper.

Finally, in the last part we shall have occasion to generalize the properties of the Laplacian from scalar functions to tensors of arbitrary rank.

Part I. Definitions

1. We shall consider a space S which is a compact orientable coördinate space of an arbitrary dimension n with an analytic positive-definite fundamental tensor $g_{ij}(x)$. In order to avoid a known trivial complication in writing, we shall exclude $n = 2$, and thus assume $n \geqq 3$.

The family of analytic functions and tensors on S, or on a specified open set in S will be denoted by C_a, and C_q, $0 \leqq q \leqq \infty$, will denote the family of functions or tensors of class q.

2. A point on S will be denoted by $x, y, z, \bar{x}, \xi, \eta, \cdots$, its coördinates by a corresponding subscript. The invariant volume element on S gives rise to a theory of Lebesgue measure and Lebesgue integration with the customary properties, except for the fact that the total measure of S is finite. The measure element will be denoted by dv, or more specifically, by dv_x, dv_y, dv_ξ, etc. Thus

$$\int_A f\,dv \quad \text{or} \quad \int_A f(x)\,dv_x$$

will denote the integral of the function f over the set A. In case $A = S$, we shall simply write

$$\int f\,dv \quad \text{or} \quad \int f(x)\,dv_x.$$

3. The class of functions $f(x)$ which are bounded and measurable on S, or on a specified subset of S, will be denoted by B. Furthermore we shall consider, on the total space S, the Hilbert space H of all (real) measurable functions of integrable square, with the inner product

$$(f, g) = \int fg\,dv.$$

We shall also use the symbol $\|f\|^2$ for (f, f). It follows easily, as in the case of a Euclidean region S, that every subspace C_q of H is dense in H. It is our main objective to prove that C_a is likewise dense in H.

[3] D. Hilbert, *Grundzüge einer allgemeinen Theorie der linearen Integralgleichungen*, 1912, chapter 18.

[4] *Problèmes mixtes et problèmes sur des variétés closes*, etc., Annales de la Société Polonaise de Mathématique, vol. 12 (1933), pp. 35–54; p. 43.

4. Functions of two variables, $f(x, y)$, $f(x, \xi)$, etc., will be defined either on the product space $S^2 = S \times S$, or on a specified part of it. It is clear what will be meant by the symbols C_a^2, C_q^2, B^2.

5. Corresponding to any point x^0 there exists a neighborhood $U = U(x^0)$ such that any two points $x, y \subset U$ have a unique geodesic distance $R \equiv R(x, y)$. The *square* of the geodesic distance will be denoted by $\Omega \equiv \Omega(x, y)$. If U is sufficiently small, then $\Omega(x, y) \subset C_a^2$. Owing to the bicompactness of S and S^2 we can find a number $\rho_0 > 0$ such that $\Omega(x, y) \subset C_a^2$ if $R(x, y) < \rho_0$.

The local analyticity of $\Omega(x, y)$ becomes obvious if $\Omega(x, y)$ is expressed in normal coördinates. The normal coördinates z_i originating at the point x_i are analytic functions

$$z_i = z_i(x, y) = y_i - x_i + (\text{higher terms in } x \text{ and } y)$$

defined for $x, y \subset U(x^0)$. They are characterized by the relations

$$\bar{g}_{ij}(z)z_j = g_{ij}(x)z_j, \tag{5}$$

in which $\bar{g}_{ij}(z)$ is the fundamental tensor expressed in the z-coördinates originating at the variable point x, with the initial values $\bar{g}_{ij}(x) = g_{ij}(x)$; the tensor $g_{ij}(x)$ being given in a fixed coördinate system of which x and y are arbitrary points. Important properties of the functions $z_i(x, y)$ are

$$\frac{\partial z_i}{\partial x_j} = -\delta_{ij}, \qquad \frac{\partial z_i}{\partial y_j} = \delta_{ij}, \qquad \text{if } z = 0, \tag{6}$$

$$\frac{\partial \bar{g}_{ij}(z)}{\partial z_k} = 0, \qquad \text{if } z = 0, \tag{7}$$

$$\Omega(x, y) = \bar{g}_{ij}(z)z_iz_j \equiv g_{ij}(x)z_iz_j.\text{[5]} \tag{8}$$

6. The family $C_{q,r}^2$, $0 \leqq r < n$, consists of functions $Q(x, y)$ with the following properties. $Q(x, y)$ is not defined for $x = y$. Everywhere else on S^2 it belongs to class C_q^2. Inside some belt $R(x, y) < \rho$, $Q(x, y)$ and each of its derivatives is representable as a quotient

$$\frac{A(x, y)}{R(x, y)^\mu}; \tag{9}$$

also, within this belt, $Q(x, y)$ has the order of magnitude $O[R(x, y)^{-r}]$, and each derivative of order $q_0(\leqq q)$ has the order of magnitude $O[R(x, y)^{-r-q_0}]$.

We shall also consider a class of functions B_r^2, $0 \leqq r < n$. The function $Q(x, y)$ belongs to B_r^2 if it is measurable on S^2, bounded outside some belt $R(x, y)$

[5] For the existence of normal coördinates with these properties it is not necessary to assume that our space is analytic. They exist if S is a space of class $p \geqq 2$, and they are themselves of class $p - 2$. Checking all steps of our argument, we could easily see that our Theorem I holds if $p = 6$, but normal coördinates can be avoided altogether and the class of the space reduced by several units. Compare G. Giraud, loc. cit.

$< \rho$, and $O[R(x, y)^{-r}]$ within this belt. For such functions we shall consider the process of convolution. For instance

$$Q' * Q'' = \int Q'(x, \xi) Q''(\xi, y) dv_\xi . \tag{10}$$

7. The symbol $\Phi_\rho(t)$ will denote any function in $0 \leqq t < \infty$ with the following properties. The function has derivatives of all orders, is non-negative throughout, has the value 1 in $0 \leqq t \leqq \frac{1}{2}\rho$, values between 0 and 1 in $\frac{1}{2}\rho < t \leqq \rho$, and the value 0 in $\rho \leqq t < \infty$.

8. The symbol $P(x, y)$ will denote any fixed function with the following properties:

$$P(x, y) = P(y, x), \tag{11}$$

$$P(x, y) \subset C_{4, n-2} , \tag{12}$$

and there exists a $\rho > 0$ such that

$$P(x, y) = \frac{\gamma}{R(x, y)^{n-2}}, \qquad \text{if } R(x, y) < \rho \tag{13}$$

for some prescribed constant γ. We shall prescribe

$$\gamma = \frac{\Gamma(\frac{1}{2}n - 1)}{4\pi^{\frac{1}{2}n}}. \tag{14}$$

The existence of such a function is trivial. For example, the function

$$P(x, y) = \gamma \frac{\Phi_\rho[R(x, y)]}{R(x, y)^{n-2}} \tag{15}$$

will do.

Any function $P(x, y)$ with these properties has the further important property

$$\Delta_x P(x, y) \subset C^2_{2, n-2} . \tag{16}$$

This property states essentially that the Laplacian, if applied to our function $P(x, y)$, does not raise the exponent of its singularity on the diagonal $x = y$; whereas, as a rule, if applied to a function from $C^2_{p,r}$, it increases the exponent r by 2. The proof of (16) follows easily if $R(x, y)$ is expressed in normal coördinates issuing from x; in such coördinates z_i we have[6]

$$\Delta_z R = \left(1 - \frac{n}{2}\right) \frac{\partial \log g(z)}{\partial z_i} z_i \cdot R^{-n},$$

and, by (7), the factor of R^{-n} on the right is $O(R^2)$.

[6] W. Feller, *Lösungen der linearen partiellen Differentialgleichungen zweiter Ordnung vom elliptischen Typus*, Mathematische Annalen, vol. 102 (1930), pp. 633–649; p. 639, footnote.

It will be important for us to consider not the Laplacian $\Delta\varphi$ itself, but the operator

$$\Delta^c\varphi = \Delta\varphi - c\varphi \tag{17}$$

for some positive constant c. The constant c being fixed, we shall write

$$K(x, y) = \Delta_x^c P(x, y), \qquad \bar{K}(x, y) = K(y, x) = \Delta_y^c P(y, x). \tag{18}$$

These functions again have the property (16).

9. The distance $R(x, y)$ is comparable in size to the Euclidean distance

$$E(x, y) = \{(y_i - x_i)^2\}^{\frac{1}{2}}. \tag{19}$$

Hence, the integral

$$\psi(y) = \int \varphi(x)Q(x, y)\, dv_x \tag{20}$$

exists for every y if $\varphi(x) \subset B$, $Q(x, y) \subset C_{0,r}$, $r < n$. For fixed $Q(x, y)$ and variable $\varphi(x)$ the integral operator (20) will be written

$$\psi = Q\varphi. \tag{21}$$

We shall be mainly interested in the operators $P\varphi$ and $K\varphi$ resulting from the functions $P(x, y)$, $K(x, y)$ that were introduced before.

Part II. Preliminaries

1. We shall state several properties of the expression (20), all of which are trivially true if $Q(x, y)$ happens to vanish inside some belt $R(x, y) < \rho_0$.

LEMMA 1. *If $\varphi(x) \subset B$, $Q(x, y) \subset C_{0,r}^2$, $r < n$, then the function* (20) *belongs to C_0^*; more precisely, the modulus of uniform continuity of $\psi(y)$ depends only on $M = \operatorname{l.u.b.}_{x \subset S} |\varphi(x)|$.*

Proof. With the function $\Phi_\rho(t)$ of Part I, §7, we form the expression $\psi_\sigma(y) = \int \varphi(x)Q(x, y)[1 - \Phi_\sigma\{R(x, y)\}]\,dv_x$. For σ fixed, our lemma holds for $\psi_\sigma(y)$ in the place of $\psi(y)$. But $\psi_\sigma(y)$ tends to $\psi(y)$ as $\sigma \to 0$, uniformly in y and M.

LEMMA 2. *If $\varphi(x) \subset B$, $Q(x, y) \subset C_{\nu,r}^2$, $r < n - \nu$, then the function* (20) *belongs to C_ν, again uniformly in M, and the partial derivatives of $\psi(y)$ may be obtained by partial differentiation under the integral. For example,*

$$\frac{\partial\psi}{\partial y_1} = \int \varphi(x)\frac{\partial Q}{\partial y_1}\, dv_x. \tag{22}$$

Proof. By using induction, it is sufficient to consider the case $\nu = 1$. Since $\dfrac{\partial Q}{\partial y_1} \subset C_{0,r+1}$, $r + 1 < n$, the integral

$$g(y) = \int \varphi(x)\frac{\partial Q}{\partial y_1}\, dv_x \tag{23}$$

represents a function in C_0. We restrict ourselves to a coördinate neighborhood $(y_1, \cdots, y_n)$ and integrate the function $g(y)$ with respect to the variable y_1 between the fixed limits α and β. Since the integrand on the right side is a Lebesgue integrable function on S^2 we can exchange, on the right side, the two processes of integration for almost all $(y_2, \cdots, y_n)$, and hence

$$\int_\alpha^\beta g(y_1, \cdots, y_n)\,dy_1 = \psi(\beta, y_2, \cdots, y_n) - \psi(\alpha, y_2, \cdots, y_n)$$

for almost all $(y_2, \cdots, y_n)$. But both sides of the last relation are continuous and, therefore, it holds throughout. Making now β variable, we find that $\psi(y) \subset C_1$, and that $g(y) = \partial\psi/\partial y_1$.

LEMMA 3. *If* $\varphi(x) \subset C_1$ *and* $Q(x, y) \subset C^2_{2,n-2}$, *the function* (20) *belongs to* C_2.

Proof. It will be sufficient to show that the function (23) belongs to C_1. Relation (6) implies

$$\frac{\partial z_i}{\partial x_j} + \frac{\partial z_i}{\partial y_j} = O[R(x, y)].$$

Hence, remembering that $Q(x, y) = A(x, y)\, R(x, y)^{-\mu}$ for $R(x, y) < \rho_0$, and using expression (8) for $\Omega(x, y)$, we easily find that

$$\frac{\partial Q}{\partial y_1} + \frac{\partial Q}{\partial x_1} \tag{24}$$

belongs to $C_{1,n-2}$. Hence, on account of Lemma 2, the non-trivial part of function (23) is the function

$$h(y) = \int \varphi(x)\, \frac{\partial Q(x, y)}{\partial x_1}\, dv_x, \tag{25}$$

and it is sufficient to show that the latter function belongs to C_1. We restrict the point y to a sufficiently small neighborhood $U(y^0)$ whose closure is contained within a sphere $R(x, y^0) < \rho$. Forming with our previous function $\Phi_\rho(t)$ the decomposition

$$\varphi(x) = \{\varphi(x)\Phi_\rho[R(x, y^0)]\} + \{\varphi(x) - \varphi(x)\Phi_\rho[R(x, y^0)]\},$$

we see immediately that the second component of $\varphi(x)$ contributes to the function (25) a component of class C_1. Hence we may assume that $\varphi(x)$ and its derivatives vanish outside some coördinate neighborhood and that y is a point of this neighborhood. Denoting the product of $\varphi(x)$ and the volume density $g^{\frac{1}{2}}$ by $\bar{\varphi}(x)$, we can therefore write

$$h(y) = \int_A \bar{\varphi}(x)\, \frac{\partial Q(x, y)}{\partial x_1}\, dx_1 \cdots dx_n,$$

where A is an n-dimensional Euclidean domain and $\bar{\varphi}(x)$ is a function of C_1 vanishing on the boundary of A. Let y be a fixed point within A, A_ρ the

difference of A and a small sphere of radius ρ around y, and B_ρ the inner boundary of A. By Green's formula,

$$\int_{A_\rho} \bar{\varphi}(x) \frac{\partial Q}{\partial x_1} dx_1 \cdots dx_n = -\int_{A_\rho} \frac{\partial \bar{\varphi}}{\partial x_1} Q \, dx_1 \cdots dx_n + \int_{B_\rho} \cos(\xi, \nu) \bar{\varphi}(\xi) Q(\xi, y) \, d\omega_\xi .$$

Since $Q(\xi, y) = O(\rho^{-n+2})$, the surface integral tends to 0 as $\rho \to 0$, and therefore

$$h(y) = -\int_A \frac{\partial \varphi}{\partial x_1} Q(x, y) \, dx_1 \cdots dx_n .$$

By Lemma 2, $h(y)$ belongs to C_1.

LEMMA 4. *If $Q(x, y) \subset B^2_r$ and $L(x, y) \subset B^2_s$, then $Q * L$ belongs to B^2_{r+s-n} if $r + s - n > 0$; to B^2_ϵ, for any $\epsilon > 0$, if $r + s - n = 0$; and to B^2_0 if $r + s - n < 0$.*

Proof. The lemma is known in case S is a bounded set in Euclidean space and $R(x, y)$ has the value (19). For our present case it follows from the fact that in every coördinate neighborhood the quantity $R(x, y)$ is majorized by (19) from above and below.

LEMMA 5. *If $Q(x, y) \subset C^2_{0,r}$, $r < n$, then, from some positive ν onward, the ν-th iterated kernel*

$$Q_\nu = Q * Q * \cdots * Q \qquad (\nu \text{ factors}) \tag{26}$$

belongs to C^2_0.

Proof. By Lemma 4 there exists a $Q_{\nu-1}(x, y)$ which is bounded. But

$$Q_\nu(x, y) = \int Q(x, \xi) Q_{\nu-1}(\xi, y) d\nu_\xi = \int Q_{\nu-1}(x, \xi) Q(\xi, y) d\nu_\xi .$$

Therefore, by Lemma 1, $Q_\nu(x, y)$ is continuous in x uniformly in y, and continuous in y uniformly in x. Hence $Q_\nu(x, y) \subset C^2_0$.

2. We shall now discuss the operator (21) for $\varphi \subset H$. To start with, the expression (20) is not defined for arbitrary elements φ in H. But it is defined for φ in B, and the subset B of H is dense in H. This suggests the possibility of extending the operator formally into a wider set of H. In fact, the operator (21) can be extended uniquely over the whole of H.

LEMMA 6. *For any $Q(x, y) \subset C^2_{0,r}$, $r < n$, the expression* (20) *defines a unique operator* (21) *which is bounded and completely continuous. Also, for any two such kernels $Q'(x, y)$, $Q''(x, y)$, the relation*

$$Q'[Q''(\varphi)] = (Q' * Q'')(\varphi) \tag{27}$$

holds.

Proof. We first assume that $Q(x, y)$ is continuous throughout. In this case it is known that $Q\varphi$ is bounded and completely continuous for φ in B. B being dense in H, the operator $Q\varphi$ has a unique extension into the whole space H,

and this extension is again completely continuous.[7] The value of $Q\varphi$ for an arbitrary element $\varphi \subset H$ can be computed in the following manner. Form any sequence of elements $\{\varphi_m\}$ from B of which φ is the limit in the strong topology of H: $\lim_{m \to \infty} \| \varphi - \varphi_m \| = 0$. The values $\psi_m = Q\varphi_m$ can be computed from (20) and they are in their turn convergent to an element ψ of H. This element is the desired value (21). Finally, relation (27) is obvious for $\varphi \subset B$ by Fubini's theorem, and for general φ it follows by the limiting process we have just described.

In the case of a general kernel $Q(x, y)$ we shall apply an important criterion.[8] An operator $Q\varphi$ is bounded and completely continuous in H, if corresponding to any $\epsilon > 0$ there exists a decomposition

$$Q\varphi = Q^\sigma\varphi + R^\sigma\varphi \tag{28}$$

of the following nature. $R^\sigma\varphi$ is bounded and completely continuous, and $Q^\sigma\varphi$ is a bounded operator whose bound is $\leqq \epsilon$. We put

$$Q^\sigma = Q(x, y)\Phi_\sigma[R(x, y)], \qquad R^\sigma = Q - Q^\sigma.$$

R^σ is continuous, and hence the operator $R^\sigma\varphi$ is bounded and completely continuous. As for the operator $\psi^\sigma = Q^\sigma\varphi$, we have, for $\varphi \subset B$,

$$\| \psi^\sigma \|^2 = \int | \psi^\sigma(y) |^2 \, dv_y = \int\int \varphi(\xi)\varphi(\eta) M^\sigma(\xi, \eta) dv_\xi dv_\eta, \tag{29}$$

where

$$M^\sigma(\xi, \eta) = \int Q^\sigma(\xi, y) Q^\sigma(y, \eta) \, dv_y. \tag{30}$$

The kernel $M^\sigma(\xi, \eta)$ has the following two properties:

$$M^\sigma(\xi, \eta) = 0, \text{ if } R(\xi, \eta) > 2\sigma; \tag{31}$$

there exists an exponent s, $0 \leqq s < n$, and a constant A such that

$$| M^\sigma(\xi, \eta) | \leqq \frac{A}{R(\xi, \eta)^s}, \qquad \text{if } R(\xi, \eta) \leqq 2\sigma. \tag{32}$$

Property (31) is an immediate consequence of the fact that $Q^\sigma(x, y)$ vanishes for $R(x, y) \geqq \sigma$; whereas property (32) is an easy consequence of Lemma 4. Putting

$$| \varphi(\xi)\varphi(\eta) | \leqq \frac{| \varphi(\xi) |^2 + | \varphi(\eta) |^2}{2}$$

in (29), we obtain

$$2 \| \psi^\sigma \|^2 \leqq \int \varphi(\xi)^2 \left[\int | M^\sigma(\xi, \eta) | \, dv_\eta \right] d_\xi + \int \varphi(\eta)^2 \left[\int | M^\sigma(\xi, \eta) | \, dv_\xi \right] dv_\eta. \tag{33}$$

[7] S. Banach, *Opérations linéaires*, 1932, p. 99.

[8] Banach, loc. cit., p. 96, Théorème 2.

But, owing to properties (31) and (32), there exists a positive quantity $\epsilon(\sigma)$, tending to 0 as σ tends to 0, such that

$$\int |M^\sigma(\xi, \eta)|\, dv_\eta \leqq \epsilon(\sigma), \qquad \int |M^\sigma(\xi, \eta)|\, dv_\xi \leqq \epsilon(\sigma).$$

Hence each of the integrals in (33) is $\leqq \epsilon(\sigma) \cdot ||\varphi||^2$, and therefore ψ^σ is a bounded operator on B whose bound is $\leqq \epsilon(\sigma)$. The extension of ψ^σ from B onto the total space H does not increase the bound,[9] and, therefore, the assumptions of the above mentioned criterion are fulfilled. This proves the first half of the lemma, whereas relation (27) can be proved as in the case of a continuous kernel $Q(x, y)$.

Part III. The Laplacian

1. If in Green's formula

$$\int \frac{1}{g^{\frac{1}{2}}} \frac{\partial}{\partial x_i} (g^{\frac{1}{2}} \lambda^i)\, dv_x = 0 \tag{34}$$

[λ^i is a contravariant vector belonging to C_1], we put

$$\lambda^i = \varphi \frac{\partial \varphi}{\partial x_i}, \qquad \text{or } \lambda^i = \varphi \frac{\partial \psi}{\partial x_i} - \psi \frac{\partial \varphi}{\partial x_i},$$

we obtain the relations

$$-\int \varphi \Delta^c \varphi\, dv = \int \left[g^{ij} \frac{\partial \varphi}{\partial x_i} \frac{\partial \varphi}{\partial x_j} + c\varphi^2 \right] dv_x \tag{35}$$

$$\int \varphi \Delta^c \psi\, dv = \int \psi \Delta^c \varphi\, dv. \tag{36}$$

From (35) we conclude

$$(\varphi, \Delta^c \varphi) \leqq -c(\varphi, \varphi), \tag{37}$$

and relation (36) reads

$$(\varphi, \Delta^c \psi) = (\Delta^c \varphi, \psi). \tag{38}$$

Hence, on the dense subspace C_2 of H, the operator $-\Delta^c \varphi$ is positive definite with a lower bound $\geqq c$ and symmetric. By a general theorem on operators in Hilbert space there exists a closed extension of the operator $\Delta^c \varphi$ for which the properties (37) and (38) remain in force.[10] It will be this closure of the Laplacian that will forthwith be denoted by $\Delta^c \varphi$; in fact, in order to simplify the writing, *we shall omit the upper index c, and thus write $\Delta \varphi$ instead of $\Delta^c \varphi$ until further notice.* The domain of $\Delta \varphi$ is a well defined set D of H, $C_2 \subset D \subset H$.

[9] Banach, loc. cit., p. 58.

[10] M. H. Stone, *Linear Transformations in Hilbert Space*, 1932, p. 49, Theorem 2.12.

Lemma 7. *The operator* $\Delta\varphi$ *satisfies the relations*

$$K\varphi - \varphi = P\Delta\varphi, \qquad \text{if } \varphi \subset D. \tag{39}$$

$$\bar{K}\varphi - \varphi = \Delta P\varphi, \qquad \text{if } P\varphi \subset D. \tag{40}$$

Proof. By the definitions given in Part I, §§8 and 9, relation (39) is equivalent to

$$\int [\varphi(x)\Delta_x P(x, y) - \Delta_x\varphi(x)P(x, y)]\, dv_x = \varphi(y). \tag{41}$$

It is sufficient to prove it for $\varphi \subset C_2$. For other φ it follows from the continuity of the operators $K\varphi$ and $P\varphi$. The left side of (41) is the limit, as $\sigma \to 0$, of

$$\int_{S_\sigma} \frac{1}{g^{\frac{1}{2}}} \frac{\partial}{\partial x_i} (g^{\frac{1}{2}}\lambda^i)\, dv_x, \tag{42}$$

where S_σ is the total space S minus the geodesic sphere of radius σ around the fixed point y, and

$$\lambda^i(x) = \varphi(x) \frac{\partial P(x, y)}{\partial x_i} - \frac{\partial\varphi}{\partial x_i} P(x, y).$$

By Green's formula, (42) has the value

$$-\int_{B_\sigma} \lambda^i(\xi) \frac{\partial R(\xi, y)}{\partial \xi_i}\, d\omega_\xi, \tag{43}$$

where B_σ is the boundary of S_σ and $d\omega_\xi$ is the invariant surface element on B_σ.[11] In order to estimate (43) we introduce normal coördinates $z_i = z_i(\xi, y)$ originating at the fixed point y. We also make a linear affine transformation by which $\bar{g}_{ij}(y) = \delta_{ij}$. This gives for the quantity $R(\xi, y)$ the value $(z_i z_i)^{\frac{1}{2}}$. Hence we obtain, by a trivial calculation,

$$\begin{aligned} \lambda^i(\xi) \frac{\partial R(\xi, y)}{\partial \xi_i} &= (2 - n)\gamma\cdot\varphi(\xi) \frac{1}{R^{n-1}} - \gamma \frac{\partial\varphi}{\partial z_i} \frac{z_i}{R^{n-1}} \\ &= (2 - n)\gamma\cdot\varphi(\xi) \frac{1}{\sigma^{n-1}} + O\left(\frac{1}{\sigma^{n-2}}\right). \end{aligned}$$

Therefore, (43) has the value

$$(n - 2)\gamma\omega_{n-1}\varphi(y) + o\left(\frac{\sigma^{n-1}}{\sigma^{n-1}}\right), \tag{44}$$

where ω_{n-1} is the $(n - 1)$-dimensional Euclidean volume of the unit-sphere $z_i z_i = 1$. Substituting from (14), we see that (44) tends to $\varphi(y)$ as $\sigma \to 0$, and this completes the proof of (39).

[11] W. Feller, loc. cit., §2; S. Bochner, *Remark on the theorem of Green*, this Journal, pp. 334–338.

Relation (40) will be proved if we show that for every element u of C_2,

$$(u, \bar{K}\varphi - \varphi - \Delta P\varphi) = 0. \tag{45}$$

Obviously

$$(u, \bar{K}\varphi) = \int u(x)\varphi(y)K(x, y)dv_x dv_y = (\varphi, Ku)$$

and

$$(u, \Delta P\varphi) = (\Delta u, P\varphi) = \int \Delta_x u(x)\cdot\varphi(y)P(y, x)dv_x dv_y$$
$$= (\varphi, P\Delta u).$$

Therefore the left side of (45) has the value

$$(\varphi, Ku - u - P\Delta u),$$

and this vanishes by (39).

2. Lemma 8. *If $Q(x, y) \subset C^2_{2,n-2}$, $f \subset C_2$, $\varphi \subset H$, the relation*

$$\varphi = Q\varphi + f \tag{46}$$

implies that $\varphi \subset C_2$.

Proof. Iterating (46), we obtain, using relation (27),

$$\varphi = Q_\nu\varphi + Q_{\nu-1}f + Q_{\nu-2}f + \cdots + f. \tag{47}$$

Obviously, the functions Qf, $Q_2f = Q(Qf)$, $\cdots$, $Q_{\nu-1}f = Q(Q_{\nu-2}f)$ are all bounded. From a sufficiently high ν onward, $Q_\nu(x, y) \subset B^2$ by Lemma 5, and hence by

$$\left|\int Q_\nu(x, y)\varphi(x)dv_x\right|^2 \leqq \int Q_\nu(x, y)^2 dv_x \int \varphi(x)^2 dv_x, \tag{48}$$

$Q_\nu\varphi$ is bounded. Thus, by (47), φ is bounded. Consequently by Lemma 2, $Q\varphi \subset C_1$, and hence by (46), $\varphi \subset C_1$. But then, by Lemma 3, $Q\varphi \subset C_2$, and therefore, again by (46), $\varphi \subset C_2$.

3. Lemma 9. *The equation*

$$(\Delta^c\varphi \equiv)\,\Delta\varphi = f \tag{49}$$

has a solution for any $f \subset H$.

Proof. A solution, if any, is unique, by (37). It is sufficient to show that (49) has a solution for $f \subset C_1$; in fact, if we show this, we prove that the operator $f = \Delta\varphi$ has an inverse

$$\varphi = Jf$$

on a dense set in H. By (37) this inverse is a bounded operator and has therefore an extension into the whole of H. If we re-invert the fully extended opera-

tor Jf we obviously obtain $\Delta\varphi$; in other words, the counter-domain of $\Delta\varphi$ is the whole space H.

Substituting (49) in (39), we are led to investigate the integral equation

$$K\varphi - \varphi = Pf. \tag{50}$$

The operator $K\varphi$ being completely continuous by Lemma 6, we can apply the generalized Fredholm theory.[12] According to this theory, the equations

$$K\varphi - \varphi = 0, \tag{51}$$

$$\bar{K}\psi - \psi = 0 \tag{52}$$

have the same finite number of independent solutions. If we denote such basic solutions by

$$\varphi_1, \cdots, \varphi_m; \qquad \psi_1, \cdots, \psi_m,$$

for a given element f the equation (50) has a solution, if, and only if,

$$(\psi_\mu, Pf) = 0 \qquad (\mu = 1, \cdots, m). \tag{53}$$

We remark once for all a fact which we shall tacitly use several times, namely, that any solution φ or ψ of (50), (51), (52) is automatically a function of C_2, and therefore a function of D. This follows from Lemma 8 [in conjunction with Lemma 3 for equation (50)]. Let φ_μ be any solution of (51). Combining (51) and (39), we obtain

$$P\Delta\varphi_\mu = 0. \tag{54}$$

Hence for $\psi_\mu = \Delta\varphi_\mu$, we obtain from (40), $\bar{K}\psi_\mu - \psi_\mu = 0$. Also, if $\varphi_1, \cdots, \varphi_m$ are linearly independent, then $\Delta\varphi_1, \cdots, \Delta\varphi_m$ are likewise linearly independent. Hence, the basic solutions of (52) can be assumed to have the form $\psi_\mu = \Delta\varphi_\mu$. Now, $(\psi_\mu, Pf) = (f, P\psi_\mu)$, and vanishes by (54). Thus equation (50) has a solution φ^* for an arbitrary element $f \subset C_1$. Comparing now relations (50) and (39), we obtain for the function $\psi^* = f - \Delta\varphi^*$ the equation $P\psi^* = 0$. By (40), ψ^* is a solution of (52) and may be written in the form $c_1\psi_1 + \cdots + c_m\psi_m$. In other words, the function $\varphi = \varphi^* + c_1\varphi_1 + \cdots + c_m\varphi_m$ is a solution of (49).

THEOREM I. *The operator $f = \Delta\varphi - c\varphi$ is for every fixed $c > 0$ the inverse of a completely continuous operator $\varphi = Jf$ in H. The solutions $\varphi \subset H$ of the equation*

$$\Delta\varphi = \lambda\varphi \tag{55}$$

(λ being any constant for which a solution exists) form a complete set of functions in H. They are all, automatically, contained in C_2.

Proof. From (50) we obtain

$$\varphi = K(Jf) - Pf \equiv Jf.$$

[12] Banach, loc. cit., pp. 159–161.

By Lemma 6, Pf is completely continuous, and so is $K(Jf)$, since Jf is at any rate bounded. Therefore, Jf is completely continuous too. Hence, Jf has a pure point spectrum and so has its inverse $\Delta^c\varphi$. Therefore, the solutions of the equation $\Delta\varphi - c\varphi = \lambda\varphi$, or, what amounts to the same, those of equation (55) form a complete basis in H. Substituting (55) in (39), we obtain $(K - \lambda P)(\varphi) = \varphi$, and therefore $\varphi \subset C_2$, by Lemma 8.

Part IV. The mapping theorem

1. The following theorem makes full use of our assumption that the underlying space S is analytic.

THEOREM II. *The solutions φ of* (55) *are all analytic. Thus there exists in the function space H a complete basis $\{\varphi_m\}$ whose elements φ_m are analytic throughout S.*

Proof. The analytic character of the functions φ_m is a purely local property and follows from the following theorem which for $n \geqq 3$ was first proved by J. Hadamard.[13]

LEMMA 10. *If the coefficients $g_{ij}(x)$ (the tensor g_{ij} being symmetric and positive-definite), $b_i(x)$, $c(x)$, $f(x)$ are all analytic in the neighborhood of a point $x^0 = (x_1^0, \cdots, x_n^0)$, then every solution $\varphi(x)$ of the equation*

$$g_{ij}\frac{\partial^2\varphi}{\partial x_i\,\partial x_j} + b_i\frac{\partial\varphi}{\partial x_i} + c\varphi = f \tag{56}$$

which belongs to C_2 is likewise analytic.

Proof. Since Hadamard does not carry out the last steps of his argument we shall give a brief summary of it. By a fundamental theorem of S. Kowalewsky equation (56) has always analytic solutions (locally). Hence it is sufficient to prove our theorem for $f(x) = 0$. We denote for the moment the operator (56) by $F\varphi$ and its formal adjoint by $F^*\varphi$, and we consider Hadamard's elementary solution $G(x, y)$ belonging to $F^*\varphi$. $G(x, y)$ is a function of the form

$$\gamma\frac{U(x, y)}{R(x, y)^{n-2}} + \log R(x, y)\cdot V(x, y); \qquad U(x, x) = 1, \tag{57}$$

the coefficients $U(x, y)$, $V(x, y)$ being analytic in all $2n$ variables $x_1, \cdots, x_n, y_1, \cdots, y_n$; and $F_x^*G(x, y) = 0$. For a sufficiently small spherical surface B our solution $\varphi(x)$ of $F_x\varphi = 0$ satisfies a relation

$$\varphi(y) = \int\left(\rho(\xi)\frac{\partial G(\xi, y)}{\partial\nu} + \sigma(\xi)G(\xi, y)\right)d\omega_\xi \tag{58}$$

for every point y interior to B; the coefficients $\rho(\xi)$ and $\sigma(\xi)$ are continuous, and we shall not need their precise value. This is a consequence of Green's formula, in which the volume integrals vanish because $\varphi(x)$ and $G(x, y)$ are solutions of $F = 0$ and $F^* = 0$, respectively. The term $\varphi(y)$ emanates from the

[13] *Lectures on Cauchy's Problem*, 1923, p. 102.

singularity (57) in just the same way as in the case of relation (39). In fact, the symmetry of the function $P(x, y)$ has not been used at all for the derivation of (39), and a logarithmic term of the form $\log R(x, y) \cdot V(x, y)$ would not alter the estimate (44) for $n \geqq 3$.

Now let A be any closed point set interior to B. For $y \subset A$, $\xi \subset B$, $R(\xi, y)$ has a positive bound from below; hence we can find a region $\bar{A}$ in the space of the *complex* variables $y_1, \cdots, y_n$ over which $R(\xi, y)$ is still analytic and has a real part which has a positive bound from below. Since $U(\xi, y)$ and $V(\xi, y)$ can also be continued into a complex y-neighborhood of the origin, it follows immediately that the function (58) is analytic in some neighborhood of the origin.

2. We again choose a fixed constant $c > 0$, and denote by $\varphi = Jf$ the inverse to $f = \Delta^c\varphi$ and by $J_\nu f$ the ν-th iterate of Jf.

LEMMA 11. *There exists an integer $\nu \geqq 1$ and a constant A such that for every $f \subset H$ the relations*

$$|J_\nu f(x)| \leqq A \cdot \|f\| \tag{59}$$

$$|\operatorname{grad} J_\nu f(x)| \leqq A \cdot \|f\| \tag{60}$$

hold.

Proof. We write $f_0 = f$, $f_{\mu+1} = Jf_\mu$, $\mu \geqq 0$. By (50) we have the relations

$$f_1 = Kf_1 - Pf_0, \qquad f_2 = Kf_2 - Pf_1, \qquad f_3 = Kf_3 - Pf_2,$$

etc., from which we deduce the relations

$$f_1 = Kf_1 - Pf_0, \qquad f_2 = KKf_2 - (KP + PK)f_1 + PPf_0,$$

$$f_3 = KKKf_3 - (KKP + KPK + PKK)f_2 + (KPP + PKP + PPK)f_1 - PPPf_0,$$

etc. In general,

$$f_\nu = Q_{\nu 0}f_\nu + Q_{\nu 1}f_{\nu-1} + \cdots + Q_{\nu\nu}f_0, \tag{61}$$

the functions $Q_{\nu\mu}(x, y)$ being "polynomials" of ν-th degree in the functions $P(x, y)$, $K(x, y)$. Since P and K both belong to B^2_{n-2}, there exists by Lemma 4 a $\nu \geqq 1$ such that the functions $Q_{\nu\mu}$ belong to B^2. We choose a ν with this property and hold it fast. Since Jf is a bounded operator, there exists an $\alpha > 0$ such that for every $f \subset H$

$$\|f_\mu(x)\| \leqq \alpha \cdot \|f\| \qquad (\mu = 0, 1, \cdots, \nu).$$

Therefore, by relations (61) and (48), there exists an $A > 0$ such that

$$|f_\nu(x)| \leqq A \cdot \|f\| \cdot$$

This proves (59). Raising the exponent ν by 1, we can also assume (for some other $A > 0$)

$$|f_{\nu-1}(x)| \leqq A \cdot \|f\|.$$

Relation (60) follows now from

$$f_\nu = Kf_\nu - Pf_{\nu-1}$$

by Lemma 2.

THEOREM III. *If $t(x)$ is any function on S belonging to C_∞, corresponding to any $\epsilon > 0$ there exists an analytic function $\psi^\epsilon(x)$ on S such that*

$$|\, t(x) - \psi^\epsilon(x) \,| \leqq \epsilon, \qquad |\, \text{grad}\, [t(x) - \psi^\epsilon(x)] \,| \leqq \epsilon. \tag{62}$$

Proof. Applying the operator Δ^c ν times to the function $t(x)$, we obtain a function $\varphi(x)$ for which $t(x) = J_\nu\varphi(x)$. By Theorem II there exists an analytic function $\varphi^\epsilon(x)$ such that $\|\, \varphi - \varphi^\epsilon \,\| \leqq \epsilon/A$. The function $\psi^\epsilon(x) = J_\nu\varphi^\epsilon(x)$ is again analytic by Lemma 10, and relations (62) hold by Lemma 11, if applied to the function $f(x) = \varphi(x) - \varphi^\epsilon(x)$.

Theorem III completes the proof of the

MAPPING THEOREM. *Any compact analytic Riemann space S_n can be mapped topologically-analytically onto the Euclidean E_{2n+1}.*

Part V. Tensors

If we use the mapping theorem it can be easily shown that not only the Hilbert space of scalars on S has an analytic basis but that the same is true for the Hilbert space H of tensors of any given rank. As an illustration we shall consider tensors of rank 3, $\varphi \equiv \varphi_{\alpha\beta\gamma}$. The inner product is, of course, defined by

$$(\varphi, \psi) = \int \varphi_{\alpha\beta\gamma}\psi^{\alpha\beta\gamma}\, dv \equiv \int \varphi^{\alpha\beta\gamma}\psi_{\alpha\beta\gamma}\, dv. \tag{63}$$

Given a tensor $\varphi_{\alpha\beta\gamma}$, we form with the functions (1) the scalars

$$A(\rho, \sigma, \tau; x) \equiv \varphi^{\alpha\beta\gamma}(x)\, \frac{\partial t_\rho}{\partial x_\alpha}\, \frac{\partial t_\sigma}{\partial x_\beta}\, \frac{\partial t_\tau}{\partial x_\gamma}$$

($\rho, \sigma, \tau = 1, 2, \cdots, 2n + 1$). Conversely, in every coördinate neighborhood on S the tensor components $\varphi^{\alpha\beta\gamma}$ can be expressed as linear combinations of the functions $A(\rho, \sigma, \tau; x)$, the coefficients being rational functions of the partial derivatives $\dfrac{\partial t_\rho}{\partial x_\alpha}$. Since the functions $A(\rho, \sigma, \tau; x)$ are limits in square mean of analytic functions, the tensor $\varphi_{\alpha\beta\gamma}$ is likewise a limit, in square mean, of analytic tensors.

Another more illuminating way of treating tensors is to set up a Laplacian for them and to investigate it as in the case of scalars. The definition of the Laplacian is

$$\Delta\varphi = g^{pq}\varphi_{\alpha\beta\gamma, p, q}.$$

Putting

$$\psi\Delta\varphi \equiv \psi^{\alpha\beta\gamma} g^{pq} \varphi_{\alpha\beta\gamma,p,q},$$
$$\psi \operatorname{grad} \varphi \equiv \psi^{\alpha\beta\gamma} g^{pq} \varphi_{\alpha\beta\gamma,p},$$
$$\nabla(\psi, \varphi) \equiv g^{pq} \psi^{\alpha\beta\gamma,p} \varphi_{\alpha\beta\gamma,q},$$

we have the formal identity

$$\psi\Delta\varphi = \operatorname{div}(\psi \operatorname{grad} \varphi) - \nabla(\psi, \varphi)$$

which leads again to (36), (37), (38). In order to have an analogue to (39) and (40), the quantity $P(x, y)$ must be a tensor of rank 3 in both the variables x and y separately. Everything goes through strictly analogously if we replace definition (15) by

$$P(x, y) = P_{\alpha\beta\gamma;\lambda\mu\nu}(x, y) = \frac{\gamma\Phi_\rho[R(x, y)]}{R(x, y)^{n-2}} \frac{\partial^2\Omega}{\partial x_\alpha \partial y_\lambda} \frac{\partial^2\Omega}{\partial x_\beta \partial y_\mu} \frac{\partial^2\Omega}{\partial x_\gamma \partial y_\nu}.$$

The analogy is immediate up to Lemma 10. This lemma is to be replaced by the following generalization.

LEMMA 12. *If in a system of equations*

$$g^{ij} \frac{\partial^2 \varphi_\mu}{\partial x_i \partial x_j} + b_{\mu\lambda i} \frac{\partial \varphi_\lambda}{\partial x_i} + c_{\mu\lambda}\varphi_\lambda = f_\mu \tag{64}$$

($i, j = 1, \cdots, n$; $\lambda, \mu = 1, \cdots, m$; n and m are arbitrary integers) all given functions are analytic, then every solution $\varphi_1, \cdots, \varphi_m$ which belongs to C_2 is also analytic.

Proof. We introduce auxiliary variables $t_1, \cdots, t_m$. Putting $\Phi(x, t) = t_\mu\varphi_\mu(x)$ and multiplying (64) by t_μ, and summing with respect to μ, we obtain

$$g^{ij}(x) \frac{\partial^2 \Phi}{\partial x_i \partial x_j} + t_\mu b_{\mu\lambda i} \frac{\partial^2 \Phi}{\partial x_i \partial t_\lambda} + t_\mu c_{\mu\lambda} \frac{\partial\Phi}{\partial t_\lambda} = t_\mu f_\mu. \tag{65}$$

Since $\frac{\partial^2 \Phi}{\partial t_\lambda \partial t_\mu} = 0$, we can add to the left side of (65) the term $\delta_{\lambda\mu} \frac{\partial^2 \Phi}{\partial t_\lambda \partial t_\mu}$. If we consider $\Phi(x, t)$ as a function in the $n + m$ variables $x_1, \cdots, x_n$; $t_1, \cdots, t_m$, relation (65) has the form (56), and Lemma 10 applies for sufficiently small values of t, since for small values of t the matrix of the coefficients of the second derivatives is again positive definite. Therefore $\Phi(x, t)$ is analytic, and so are the functions

$$\varphi_\mu(x) = \frac{\partial\Phi}{\partial t_\mu}.$$

PRINCETON UNIVERSITY.

POLAR CORRESPONDENCE WITH RESPECT TO A CONVEX REGION

By Fritz John

We denote as polar correspondence (abbreviated P.C.) with respect to a convex region R in projective n-dimensional space π_n any one-to-one correspondence of the points of R and the hyperplanes outside R. The study o such a correspondence is essentially the study of a contragredient vector with special consideration of the convexity of the domain of definition. In §1 of this paper the representation of a P.C. in homogeneous coördinates is discussed. A P.C. is called *positive* if a point and its polar plane are not separated by any other point and its polar plane. In §2 it is proved that every positive P.C. is continuous. §§3–4 deal with *symmetric* P.C.'s; a P.C. is called symmetric if in the neighbourhood of every point P it is approximated by an ordinary polar correspondence with respect to a quadric, denoted as the *tangential* quadric in P. A positive symmetric P.C. may be generated by a convex hypersurface in $(n + 1)$-dimensional space π_{n+1} in such a way that the line joining any point Q of the surface to a fixed point of π_{n+1} and the tangential plane in Q intersect π_n in a point and its polar respectively.

In the remainder of the paper a general class of P.C.'s is discussed, which are generated by continuous positive mass distributions on R. Given any hyperplane p outside R, the pole of p shall be that point which becomes the center of mass of R with the given mass distribution, in case p is chosen as plane at infinity. In §4 it is proved that a P.C. generated in this way is always positive and symmetric. In §§5–6 it is shown that the tangential quadric at a point P is identical with Legendre's ellipsoid of inertia of R if the polar of P is plane at infinity. Moreover, some inequalities involving R and its tangential quadrics are given.[1]

1. Let R be an *open* convex region in projective n-dimensional space π_n; i.e., an open set with the following properties:

(1) If P and Q are points of R, one of the two straight line segments bounded by P and Q belongs to R;

(2) If S denotes the set of points which are neither points of R nor boundary points of R, there is at least one hyperplane in S.

Definition. A one-to-one correspondence between the points of R and

Received October 23, 1936.

[1] I am indebted to the referee for pointing out that the methods used in §1 are closely related to those used by Steinitz in his paper *Bedingt konvergente Reihen und konvexe Systeme* in Crelle's Journal; cf. in particular vol. 146, p. 32 et seq., where Steinitz deals with convex regions in projective space. Our sets ρ and $\bar{\rho}$ appear there as number sets A and $-A$, and theorems corresponding to our Theorems 1.7 and 1.8 are given.

the hyperplanes of S is called a *polar correspondence* (abbreviated P.C.) with respect to R. If a point P and a hyperplane p are corresponding elements, then P will be called the *pole* of p and p the *polar* of P.

Let π_n be referred to homogeneous coördinates $x_1, \cdots, x_{n+1}$. Let ρ denote the set of ordered sets $x = (x_1, x_2, \cdots, x_{n+1})$ corresponding to points of R, and let σ denote the set of ordered sets $u = (u_1, u_2, \cdots, u_{n+1})$ corresponding to hyperplanes $u_\alpha x_\alpha = 0$ contained in S.[2]

For every $x \subset \rho$ our P.C. uniquely determines, but for a common factor, a $u \subset \sigma$; we assume in what follows that this factor is chosen in such a way that for a point x and its polar u the relation

$$u_\alpha x_\alpha = 1 \tag{A}$$

holds. This is possible, as obviously $u_\alpha x_\alpha \neq 0$. The u_α then become uniquely determined functions $u_\alpha = u_\alpha(x)$ of $x = (x_1, \cdots, x_{n+1})$ for $x \subset \rho$ in the given P.C.

The following properties of the u_α follow immediately from our assumptions:

1.1. *$u_\alpha(x_1, \cdots, x_{n+1})$ is homogeneous of degree -1.*

1.2. *For every $u \subset \sigma$ the equations $u_\alpha(x) = u_\alpha$ $(\alpha = 1, \cdots, n + 1)$ have a unique solution $x \subset \rho$.*

1.3. *If a new coördinate system is introduced by $x'_\alpha = a_{\alpha\beta} x_\beta$ the u_β undergo the contragredient transformation $u_\beta = a_{\alpha\beta} u'_\alpha$.*

1.4. *If $x \subset \rho$ and $y \subset \rho$, then $u_\alpha(x) y_\alpha \neq 0$.*

The convexity of R finds its expression in the following theorems.

1.5. *If $x \subset \rho$, $y \subset \rho$, the equation*

$$u_\alpha(z)(\lambda x_\alpha + \mu y_\alpha) = 0$$

has either for no z of ρ or for every z of ρ a solution with $\lambda > 0$, $\mu > 0$.

Proof. $\lambda x + \mu y$ gives for $\lambda > 0$, $\mu > 0$ one of the two segments bounded by x and y; according as this segment is contained in R or not it will be intersected by no or by every hyperplane contained in S.

1.6. *For $x, y, z \subset \rho$,*

$$u_\alpha(z)x_\alpha \cdot u_\beta(x)y_\beta \cdot u_\gamma(y)z_\gamma > 0; \tag{1}$$

in particular for $y = z$,

$$u_\alpha(y)x_\alpha \cdot u_\beta(x)y_\beta > 0.$$

Proof. Let x and y be such that

$$u_\alpha(z)(\lambda x_\alpha + \mu y_\alpha) \neq 0$$

for all $z \subset \rho$ and $\lambda > 0$, $\mu > 0$. As $u_\alpha(z)(\lambda x_\alpha + \mu y_\alpha)$ represents a continuous function of λ and μ, it will have the same sign for $\lambda = 0$, $\mu = 1$ as for $\lambda = 1$, $\mu = 0$. Thus

$$u_\alpha(z)x_\alpha \cdot u_\beta(z)y_\beta > 0 \tag{2}$$

[2] Throughout this paper the summation convention is used, the letters $\alpha, \beta, \gamma, \cdots$ ranging over $1, \cdots, n + 1$, the letters $i, k, l, \cdots$ ranging over $1, \cdots, n$.

for all $z \subset \rho$. In particular, for $z = x$ or $z = y$,

$$u_\alpha(x)y_\alpha > 0, \qquad u_\beta(y)x_\beta > 0; \tag{3}$$

hence

$$u_\alpha(x)y_\alpha \cdot u_\beta(y)x_\beta > 0. \tag{4}$$

If, on the other hand, x and y are such that

$$u_\alpha(z)(\lambda x_\alpha + \mu y_\alpha) = 0$$

for every $z \subset \rho$ and some $\lambda > 0$, $\mu > 0$, we may conclude similarly that for every $z \subset \rho$

$$u_\alpha(z)x_\alpha \cdot u_\beta(z)y_\beta < 0 \tag{2'}$$

$$u_\alpha(x)y_\alpha < 0, \qquad u_\beta(y)x_\beta < 0, \tag{3'}$$

so that (4) holds. As (4) is proved for every two elements of ρ, we also have

$$u_\beta(z)y_\beta \cdot u_\gamma(y)z_\gamma > 0.$$

Thus from (2), (3), (4) or from (2′), (3′), (4) follows (1).

DEFINITION. Let x^0 be a fixed element of ρ. By $\bar{\rho}$ we denote the subset of all elements y of ρ for which

$$u_\alpha(x^0)y_\alpha > 0.$$

Obviously for every $y \subset \rho$ either $y \subset \bar{\rho}$ or $-y \subset \bar{\rho}$. It follows immediately from 1.6, that

1.7. *If $x \subset \bar{\rho}$ and $y \subset \bar{\rho}$, then*

$$u_\alpha(x)y_\alpha > 0.$$

Thus $\bar{\rho}$ may just as well be generated by any other one of its elements as by x^0.

1.8. *If $x \subset \bar{\rho}$ and $y \subset \bar{\rho}$, then $\lambda x + \mu y \subset \bar{\rho}$ for $\lambda > 0$, $\mu > 0$.*

1.9. *Let $x^\nu \subset \bar{\rho}$ for $\nu = 1, 2, \cdots$ and $\lim_{\nu\to\infty} x^\nu = x \subset \rho$.*

Then there exists a subsequence of the $u(x^\nu)$ which converges to some $(u) \neq 0$.

Proof. There is certainly a subsequence of the $u(x^\nu)$ for which

$$v^\nu = u(x^\nu)\,[\sum_\alpha u_\alpha^2(x^\nu)]^{-\frac{1}{2}}$$

converges toward some $v \neq 0$. As $u_\alpha(x^\nu)x_\alpha^\nu = 1$ (ν not summed), we have for this subsequence

$$v_\alpha^\nu x_\alpha^\nu = [\sum_\alpha u_\alpha^2(x^\nu)]^{-\frac{1}{2}}, \qquad v_\alpha x_\alpha = \lim_{\nu\to\infty} [\sum_\alpha u_\alpha^2(x^\nu)]^{-\frac{1}{2}}.$$

The limit on the right side must be $\neq 0$, for the plane $v_\alpha y_\alpha = 0$ cannot contain an interior point of R, as it is the limit of exterior planes of R. Thus $[\sum u_\alpha^2(x^\nu)]^{\frac{1}{2}}$ converges for that subsequence towards a limit $\neq 0$, and therefore the corresponding $u(x^\nu)$ converge toward some $u \neq 0$.

1.10. *$\bar{\rho}$ is an open set.*

Proof. If $x \subset \bar{\rho}$ and $x = \lim_{\nu\to\infty} x^\nu$, then $\lim_{\nu\to\infty} u_\alpha(x)x^\nu_\alpha = u_\alpha(x)x_\alpha = 1$. Thus $u_\alpha(x)x^\nu_\alpha > 0$ for sufficiently large ν; i.e., $x^\nu \subset \bar{\rho}$ for sufficiently large ν.

2. Definition. The P.C. is called *positive* if for any $x \subset \rho$, $y \subset \rho$ and $x \neq y$

$$u_\alpha(x)y_\alpha \cdot u_\beta(y)x_\beta > 1. \tag{6}$$

(Compare with 1.6.)

Let P and Q be the points of R corresponding to x and y; let P' and Q' be the points of intersection of PQ with the polars of P and Q, respectively. Then the cross ratio (PP'/QQ') is given by

$$1 - \frac{1}{u_\alpha(y)x_\alpha \cdot u_\beta(x)y_\beta}.$$

Thus (6) means that the pair of points PP' is not separated by the pair of points QQ'.

2.1. *In a positive P.C. the functions $u_\alpha(x_1, \cdots, x_{n+1})$ are continuous.*

Proof. Let $x^\nu \subset \rho$ for $\nu = 1, 2, \cdots$ and let $\lim_{\nu\to\infty} x^\nu = x \subset \rho$. We have to prove that $\lim_{\nu\to\infty} u(x^\nu) = u(x)$. According to 1.9 a subsequence of the $u(x^\nu)$ will converge towards some $v \neq 0$. We restrict ourselves to the consideration of that subsequence, which may again be denoted by $u(x^\nu)$. Our statement will be proved if $v = u(x)$. Let us assume that $\lim_{\nu\to\infty} u(x^\nu) = v \neq u(x)$. There are two possible cases: (a) $v \subset \sigma$, (b) v is on the boundary of σ; i.e., corresponds to a plane of support of R.

We deal first with the case (a), where $\lim_{\nu\to\infty} u(x^\nu) = v \subset \sigma$. There is a $y \subset \rho$ such that $v = u(y)$, and $y \neq x$, since $u(y) \neq u(x)$. Without restriction of generality we may assume that $x \subset \bar{\rho}$. Then according to 1.10 all but a finite number of the x^ν, which may be neglected, are contained in $\bar{\rho}$. From $\lim_{\nu\to\infty} x^\nu = x$, $\lim_{\nu\to\infty} u(x^\nu) = u(y)$, and $u_\alpha(x^\nu)x^\nu_\alpha = 1$, it follows that

$$u_\alpha(y)x_\alpha = 1;$$

therefore $y \subset \bar{\rho}$. Moreover, for any $z \subset \rho$,

$$u_\alpha(z)x^\nu_\alpha \cdot u_\beta(x^\nu)z_\beta \geqq 1;$$

thus for $\nu \to \infty$,

$$u_\alpha(z)x_\alpha \cdot u_\beta(y)z_\beta \geqq 1.$$

Let $z = \mu x + \lambda y$, where $\lambda > 0$, $\mu > 0$. Then $z \subset \bar{\rho}$ and

$$\begin{aligned} 1 &\leqq [u_\alpha(\lambda y + \mu x)x_\alpha][u_\beta(y)(\lambda y + \mu x_\beta)] \\ &= [u_\alpha(\mu x + \lambda y)x_\alpha](\mu + \lambda). \end{aligned} \tag{7}$$

On the other hand,

$$u_\alpha(\mu x + \lambda y)x_\alpha = \frac{1}{\mu}\,u_\alpha(\mu x + \lambda y)(\mu x_\alpha + \lambda y_\alpha) - \frac{\lambda}{\mu}\,u_\alpha(\mu x + \lambda y)y_\alpha$$

$$= \frac{1}{\mu} - \frac{\lambda}{\mu}\,u_\alpha(\mu x + \lambda y)y_\alpha$$

$$< \frac{1}{\mu} - \frac{\lambda}{\mu}\cdot\frac{1}{u_\alpha(y)(\mu x_\alpha + \lambda y_\alpha)},$$

by use of (6) and the fact that $u_\alpha(y)(\mu x_\alpha + \lambda y_\alpha) > 0$, since z and $y \subset \bar{\rho}$. The right member is

$$\frac{1}{\mu} - \frac{\lambda}{\mu}\cdot\frac{1}{\mu + \lambda} = \frac{1}{\mu + \lambda}.$$

But the last result contradicts (7).

In case (b), we have $\lim_{\nu\to\infty} u(x^\nu) = v$, where $v_\alpha y_\alpha = 0$ is the equation of a plane of support of R (or of ρ). Then there exists a y on the boundary of ρ such that $v_\alpha y_\alpha = 0$; y may be chosen in such a way that $u_\alpha(x)y_\alpha > 0$. Then for $\lambda > 0$, $\mu > 0$,

$$u_\alpha(x)(\mu x_\alpha + \lambda y_\alpha) = \mu + \lambda u_\alpha(x)y_\alpha > 0,$$

i.e., $\mu x + \lambda y \subset \bar{\rho}$ for all positive μ, λ. Then

$$u_\alpha(\mu x + \lambda y)(\mu_1 x_\alpha + \lambda_1 y_\alpha) > 0,$$

if μ_1 and λ_1 are positive. Thus for $\mu_1 \to 0$ we have

$$u_\alpha(\mu x + \lambda y)y_\alpha > 0, \tag{8}$$

since $u_\alpha(\mu x + \lambda y)y_\alpha = 0$ is not possible. On the other hand, it follows from

$$u_\alpha(\mu x + \lambda y)x_\alpha^\nu \cdot u_\beta(x^\nu)(\mu x_\beta + \lambda y_\beta) \geqq 1$$

for $\nu \to \infty$ that

$$u_\alpha(\mu x + \lambda y)x_\alpha \cdot (\mu x_\beta + \lambda y_\beta)v_\beta \geqq 1.$$

As $v_\alpha x_\alpha = 1$, $v_\alpha y_\alpha = 0$, we have

$$u_\alpha(\mu x + \lambda y)\mu x_\alpha \geqq 1;$$

subtracting this from $u_\alpha(\mu x + \lambda y)(\mu x_\alpha + \lambda y_\alpha) = 1$, we obtain

$$u_\alpha(\mu x + \lambda y)y_\alpha \leqq 0,$$

in contradiction to (8).

3. Let the P.C. be positive. Let

$$U(x, y) = \int_0^1 u_\alpha[(1 - t)x + ty](y_\alpha - x_\alpha)dt. \tag{9}$$

$U(x, y)$ is defined for $x \subset \bar{\rho}$ and $y \subset \bar{\rho}$, as then also $(1 - t)x + ty \subset \bar{\rho}$ for $0 \leqq t \leqq 1$, according to 1.8.

3.1. $U(x, y) = -U(y, x), \qquad U(x, \lambda y) = U\left(\dfrac{x}{\lambda}, y\right).$

DEFINITION. A positive P.C. is called *symmetric*, if for all $x, y, z \subset \bar{\rho}$

$$U(x, y) + U(y, z) + U(z, x) = 0. \tag{10}$$

3.2. *In a positive symmetric P.C.*

$$\frac{\partial U(x, y)}{\partial y_\alpha} = u_\alpha(y).$$

Proof. From (10) we obtain for $z_\beta = y_\beta + \delta_\alpha^\beta h$

$$U(x, z) - U(x, y) = U(y, z) = \int_0^1 u_\alpha[y + (z - y)t]\,hdt$$

or

$$\lim_{h \to 0} \frac{U(x, z) - U(x, y)}{h} = \lim \int_0^1 u_\alpha[y + (z - y)t]\,dt = u_\alpha(y).$$

3.3. *A positive P.C. is symmetric if there exists a function* $u(x_1, \cdots, x_{n+1}) > 0$ *defined in* $\bar{\rho}$ *such that*

$$\frac{\partial \log u}{\partial x_\alpha} = u_\alpha. \tag{11}$$

Proof. If there is a function u satisfying (11), then

$$U(x, y) = \log \frac{u(y)}{u(x)}, \tag{12}$$

and (10) will be obviously satisfied. If on the other hand (10) holds, the function $u(y)$ defined by

$$u(y) = e^{U(x,y)}u(x) \qquad (u(x) > 0), \tag{13}$$

where x is some point of $\bar{\rho}$, will satisfy (11) according to 3.2, and will be positive.

It follows from (12) that all solutions u of (11) can only differ by a constant factor and are given by (13).

3.4. *The function $u(x)$ is homogeneous of degree one.*

Proof. Let u be given by (13). According to (10) for $\lambda > 0$,

$$u(\lambda y) = e^{U(x,\lambda y)}u(x) = e^{U(x,y)+U(y,\lambda y)}u(x) = u(y)e^{U(y,\lambda y)}.$$

Now from 1.1 and (9) it is easily deduced that $U(y, \lambda y) = \log \lambda$.

Let π_{n+1} denote the projective $x_1 \cdots x_{n+1}$ y-space and let Y be the point $x_1 = x_2 = \cdots = x_{n+1} = 0$ and π_n the plane $y = 0$. Then the equation $y = u(x_1, \cdots, x_{n+1})$ for $x \subset \bar{\rho}$ defines, according to 3.4, a certain surface Σ in π_{n+1}. As u has continuous derivatives, Σ has a continuously changing tangent plane

at every point. Moreover, for every point Q of Σ the line YQ intersects $y = 0$ in a point of R.

3.5. *If Q is a point of Σ, then the point of intersection of the line YQ with π_n and the $(n - 1)$-flat of intersection of the tangent plane in Q with π_n are respectively pole and polar in our P.C.*

Proof. The intersection of the tangent plane in Q with $y = 0$ is given by

$$y_\alpha \frac{\partial u}{\partial x_\alpha} = 0,$$

or according to (11) by $u_\alpha(x)y_\alpha = 0$.

3.6. *Σ is a convex surface with only regular points and planes of support.*[3]

Proof. Let P_1 and $P_2 \subset R$. The 2-flat through P_1P_2Y intersects Σ along some curve L. From the geometrical interpretation of the positiveness of the P.C. (cf. the definition, §2) it follows that L is a convex curve; for if P_1', P_2' are the points of intersection of the line P_1P_2 with the tangents of L in the points of intersection of P_1Y and P_2Y with L, then P_1P_1' and P_2P_2' do not separate each other. Σ is a convex surface, as it is intersected in a convex curve by every 2-flat through Y. The regularity of the points and planes of support of Σ follows immediately from the one-to-one character of the P.C.

The following inverse statement is easily proved:

3.7. *Any surface Σ in π_{n+1} will generate a positive, symmetric P.C. with respect to the convex region R in π_n in the manner of* 3.5, *if*

(a) *for every point Q of Σ the line YQ intersects π_n in a point of R;*

(b) *Σ together with R and the boundary of R forms the boundary of a convex region in π_{n+1};*

(c) *Σ has no points in common with π_n;*

(d) *Σ has only regular points and planes of support.*

4. Throughout this section we assume that the P.C. is positive and that the functions $u_\alpha(x)$ have continuous first derivatives. We put

$$\frac{\partial u_\alpha}{\partial x_\beta} = u_{\alpha\beta}(x).$$

Definition. The correlation in which a point ξ corresponds to the hyperplane given by

$$[u_{\alpha\beta}(x) + 2u_\alpha(x)u_\beta(x)]\,\xi_\alpha\eta_\beta = 0 \tag{14}$$

will be denoted as the tangential correlation at x. If we write

$$v_\alpha(\xi) = \frac{[u_{\beta\alpha}(x) + 2u_\alpha(x)u_\beta(x)]\,\xi_\beta}{[u_{\gamma\beta}(x) + 2u_\gamma(x)u_\beta(x)]\,\xi_\gamma\xi_\beta}, \tag{15}$$

[3] Regular points are points through which only one plane of support passes; regular planes of support are planes containing only one point of contact. Cf. the definition in Bonnesen and Fenchel, *Konvexe Körper*, pp. 13–15.

the plane $v_\alpha(\xi)\eta_\alpha = 0$ corresponds to the point ξ in the tangential correlation and $v_\alpha(\xi)\xi_\alpha = 1$.

4.1. (a) $v_\alpha(x) = u_\alpha(x)$, *i.e., the polar of x is the same for the P.C. and the tangential correlation at* (x);

(b) $v_\alpha(\xi)x_\alpha = 0$ *implies* $0 = v_\alpha(x)\xi_\alpha$, *i.e., in the tangential correlation the poles of planes through* (x) *lie on the polar of* (x);

(c) $v_{\alpha\beta}(x) = \left(\dfrac{\partial v_\alpha(\xi)}{\partial \xi_\beta}\right)_{\xi=x} = u_{\alpha\beta}(x)$, *i.e. the tangential correlation gives the best approximation to the P.C. in the neighbourhood of x.*

Proof. From 1.1 and (A) it follows that

$$u_{\alpha\beta}(x)x_\alpha = u_{\beta\alpha}(x)x_\alpha = -u_\beta(x), \qquad u_{\alpha\beta}(x)x_\alpha x_\beta = -u_\beta(x)x_\beta = -1.$$

4.2. *When u_α has continuous derivatives, the P.C. is symmetrical, if and only if* $u_{\alpha\beta}(x) = u_{\beta\alpha}(x)$.

Proof. $U(x, y)$ is the line integral of the function u over the straight line segment joining the points x and y of $\bar{p}$.

Since for $u_{\alpha\beta} = u_{\beta\alpha}$ the tangential correlation (14) is the polar correspondence with respect to a quadric, we see that a P.C. is symmetric if in the neighborhood of a point it can be approximated by an ordinary polar correspondence with respect to a quadric. This quadric, which is given by

$$[u_{\alpha\beta}(x) + 2u_\alpha(x)u_\beta(x)]\xi_\alpha\xi_\beta = 0,$$

will be called the *tangential quadric* at x. It follows from 4.1 that

4.3. *If Σ is the surface in π_{n+1} generating a symmetric P.C., the tangential quadric at a point P of R is the intersection of π_n with that $(n + 1)$-dimensional quadric which* (a) *has a contact of the second order with Σ at the point of intersection of PY with Σ and* (b) *for which Y is the pole of the plane π_n.*

4.4. *If the tangential quadric at x in a positive symmetric P.C. is non-degenerate, it has no point in common with the polar of x. (This implies that the tangential quadric is projectively equivalent to an ellipsoid.)*

Proof. According to (6),

$$u_\alpha(\xi)x_\alpha \cdot u_\beta(x)\xi_\beta \geqq 1;$$

in particular for $\xi = x + dx$, if we neglect terms of higher than the second order in dx,

$$\begin{aligned} 1 &\leqq [u_\alpha(x + dx)x_\alpha][u_\beta(x)(x_\beta + dx_\beta)] \\ &= [1 - u_\alpha(x + dx)dx_\alpha][1 + u_\beta(x)dx_\beta] \\ &= [1 - u_\alpha(x)dx_\alpha - u_{\alpha\gamma}(x)dx_\alpha dx_\gamma + \cdots][1 + u_\beta(x)dx_\beta] \\ &= 1 - [u_\alpha(x)dx_\alpha]^2 - u_{\alpha\gamma}(x)dx_\alpha dx_\gamma + \cdots. \end{aligned}$$

Consequently for any $dx = \eta$,

$$[u_{\alpha\beta}(x) + 2u_\alpha(x)u_\beta(x)]\eta_\alpha\eta_\beta \leqq [u_\alpha(x)\eta_\alpha]^2.$$

From this inequality it follows that, provided the quadric

$$[u_{\alpha\beta}(x) + 2u_\alpha(x)u_\beta(x)]\,\eta_\alpha\eta_\beta = 0$$

is non-degenerate, it can have no points in common with the plane $u_\alpha(x)\eta_\alpha = 0$.

5. Let the convex region R be covered with mass of density μ. We assume that μ is a continuous *positive* function in R and on its boundary, and that μ is invariant under collineations.

DEFINITION. Let p be any hyperplane of S. If p is taken as plane at infinity in a non-homogeneous coördinate system, R will have a certain center of mass P under the given mass distribution. P will be called the *pole of mass* of p and p the *polar of mass* of P.

If an ellipsoid or a simplex is covered with homogeneous mass, the pole of mass of a plane coincides with the pole in the ordinary definition.

As a consequence of the well-known theorem that the center of mass of a convex region with a positive mass distribution lies in the interior of the region, and is invariant under affine transformations, we have:

5.1. *Every plane of S has a uniquely determined pole of mass, which is contained in R.*

In this and the following paragraphs we shall always refer R to a *non-homogeneous* coördinate system $x_1, \cdots, x_n$, the plane at infinity being a plane of S and the origin a point of R. Let $q_i x_i = 1$ be the equation of a plane p of S. In order to calculate the pole of mass P of p, we introduce new non-homogeneous coördinates

$$x'_i = \frac{x_i}{1 - q_k x_k} \qquad (i = 1, \cdots, n)$$

such that p becomes plane at infinity in x'-space. If $\xi'_1, \cdots, \xi'_n$ denote the coördinates of the center of mass of R in x'-space referred to x' coördinates and $\xi_1, \cdots, \xi_n$ the coördinates of the same point referred to the x-system, we have

$$\xi'_i = \frac{\int_{R'} \mu x'_i \, dx'_1 \cdots dx'_n}{\int_{R'} \mu \, dx'_1 \cdots dx'_n},$$

or in the x-system,

$$\frac{\xi_i}{1 - q_k \xi_k} = \frac{\int_R \mu \dfrac{x_i}{1 - q_k x_k} J \, dx_1 \cdots dx_n}{\int_R \mu J \, dx_1 \cdots dx_n},$$

when

$$J = \frac{\partial(x'_1, \cdots, x'_n)}{\partial(x_1, \cdots, x_n)} = \frac{1}{(1 - q_l x_l)^{2n}} \,|\, \delta_i^k(1 - q_m x_m) + x_i q_k \,|$$
$$= (1 - q_l x_l)^{-n-1}.$$

Thus

$$\frac{\xi_i}{1 - q_k\xi_k} = \frac{\int_R \mu x_i(1 - q_l x_l)^{-n-2} dx_1 \cdots dx_n}{\int_R \mu (1 - q_l x_l)^{-n-1} dx_1 \cdots dx_n}. \tag{16}$$

Let us write

$$F(q_1, \cdots, q_n) = \left(\int_R \mu(1 - q_l x_l)^{-n-1} dx_1 \cdots dx_n\right)^{-\frac{1}{n+1}}. \tag{17}$$

If F_i stands for $\partial F/\partial q_i$, (16) may be written

$$\frac{\xi_i}{1 - q_k\xi_k} = -\frac{F_i}{F} \qquad (i = 1, \cdots, n). \tag{18}$$

Solving for ξ_i, we finally obtain

$$\xi_i = \frac{-F_i}{F - q_k F_k}. \tag{19}$$

$F(q_1, \cdots, q_n)$ is defined for all $(q_1, \cdots, q_n)$ for which $q_k x_k < 1$ for all $(x_1, \cdots, x_n) \subset R$; i.e., for all q contained in the polar region R' of R with respect to the unit sphere. F gives essentially the mass of R in any coördinate system.

5.2. *$F(q) = F(q_1, \cdots, q_n)$ is strictly concave for $q \subset R'$.*

Proof. We have to prove that for $0 < \vartheta < 1$ and $q' \neq q''$

$$F[\vartheta q' + (1 - \vartheta)q''] > \vartheta F(q') + (1 - \vartheta)F(q'').$$

Let

$$a(x) = \frac{1 - q_l' x_l}{\mu^{\frac{1}{n+1}}}, \qquad b(x) = \frac{1 - q_l'' x_l}{\mu^{\frac{1}{n+1}}},$$

$$c(x) = \frac{1 - (\vartheta q_l' + (1 - \vartheta)q_l'')x_l}{\mu^{\frac{1}{n+1}}}.$$

Then $c(x) = \vartheta a(x) + (1 - \vartheta)b(x)$ and

$$\begin{aligned} F[\vartheta q' + (1 - \vartheta)q''] &= \left[\int_R c^{-(n+1)}(x)\, dx_1 \cdots dx_n\right]^{-\frac{1}{n+1}} \\ &= \left\{\int [\vartheta a + (1 - \vartheta)b]^{-(n+1)} dx_1 \cdots dx_n\right\}^{-\frac{1}{n+1}} \\ &> \left\{\int [\vartheta a]^{-(n+1)} dx_1 \cdots dx_n\right\}^{-\frac{1}{n+1}} + \left\{\int [(1 - \vartheta)b]^{-(n+1)} dx_1 \cdots dx_n\right\}^{-\frac{1}{n+1}} \\ &= \vartheta F(q') + (1 - \vartheta)F(q'') \end{aligned}$$

according to Minkowski's inequality.[4]

[4] Cf. Hardy, Littlewood, Pólya, *Inequalities*, p. 146, Theorem 198.

Obviously the equation $F = F(q_1, \cdots, q_n)$ will be the equation of an open, concave surface Σ' in $q_1 \cdots q_n F$-space lying above the region R'. Every parallel to the F-axis through a point of R' will intersect Σ' in a single point. As F is *strictly* concave, only one point of Σ' will be on every plane of support of Σ'. As F has continuous derivatives, there will pass only one plane of support through every point of Σ'. As $\mu > 0$ on the boundary of R, F will tend towards 0, if $(q_1, \cdots, q_n)$ approaches the boundary of R'. Thus Σ' together with R' and its boundary will form the boundary of a convex region in $q_1 \cdots q_n$ F-space. Hence through every $(n - 1)$-flat of $F = 0$ outside R' will pass one plane of support of Σ' having a point of contact on Σ' (not on its boundary).

Let Σ be the polar reciprocal of Σ', i.e., the surface consisting of the poles of the planes of support of Σ' with respect to the unit sphere.[5] Σ will be given in tangential coördinates by the equation

$$q_i \xi_i + \eta F(q_1, \cdots, q_n) = 1 \qquad (q \subset R'), \tag{20}$$

$\xi_1, \cdots, \xi_n, \eta$ denoting the coördinates of a point of a tangential plane of Σ. By dualizing the previous statements about Σ', we see that Σ will be a concave surface with only regular points and planes of support; through every $(n - 1)$-flat outside R will pass exactly one plane of support of Σ having one point in common with Σ. Every parallel to the η-axis through a point of R will intersect Σ, i.e., Σ lies above R. If $q_i \xi_i = 1$ is any $(n - 1)$-flat p in $\eta = 0$ outside R, the coördinates of the point of contact of the plane of support of Σ through p are given by (20) and the equations

$$\xi_k + \eta F_k = 0 \qquad (k = 1, \cdots, n). \tag{21}$$

From (20) and (21) we find (19).

Accordingly, the projection of the point of contact is the pole of mass of p. Hence the correspondence between poles of mass and polars of mass is generated by the surface Σ in the manner of 3.5, if the point at infinity of the η-axis is taken as Y. Thus according to 3.7:

5.3. *The correspondence between poles of mass and planes of mass is a positive symmetric P.C.*

Incidentally we have proved that every point of R is pole of one hyperplane in S; in other words, that for every point P of R there is a collineation such that the image of P becomes center of mass of the transformed region. This might have been more directly concluded from (19) and 5.2 in the following way. It is sufficient to prove that the origin $\xi_1 = \xi_2 = \cdots = \xi_n = 0$, which was an arbitrary point of R, is pole of mass of some plane $q_l x_l = 1$. Thus according to (19) we have to prove that the equations $F_i = 0$ $(i = 1, \cdots, n)$ have a solution, i.e., that F is stationary for some $q \subset R'$. Now, as F is a concave positive function in R' and $F = 0$ on the boundary of R', F will have a maximum in R'; this maximum will be attained at only one point of R', as F is *strictly* concave. Thus the origin will be pole of mass of exactly one plane.

[5] Cf. Bonnesen and Fenchel, loc. cit., p. 28.

Let us calculate the tangential quadric of the P.C. (19) at a point P of R with coördinates $(x_1^0, \cdots, x_n^0)$. Let $q_i^0 y_i = 1$ be the equation of the polar p of P. The tangential quadric Q at P is according to §4 characterized by the conditions that the coördinates x_i of a point and their first derivatives x_{ik} with respect to the coördinates q_k of its polar plane have for $q = q^0$ the same values under the polar correspondence (19) and under the polar correspondence with respect to Q. In order to simplify the calculation we assume that P is origin and p plane at infinity of our non-homogeneous coördinate system. Then $x_i^0 = 0$, $q_k^0 = 0$. Q is given in tangential coördinates by

$$x_{ik} q_i q_k = 1. \tag{22}$$

For the pole $(x_1, \cdots, x_n)$ of a plane $q_i y_i = 1$ with respect to the quadric (22) is given by

$$x_i = x_{ik} q_k ;$$

hence for $q_1 = q_2 = \cdots = q_n = 0$

$$x_i = 0, \qquad \frac{\partial x_i}{\partial q_k} = x_{ik}.$$

According to (19) we have for $q_1 = q_2 = \cdots = q_n = 0$ and $x_1 = \cdots = x_n = 0$

$$F_i = 0, \qquad x_{ik} = \frac{\partial x_i}{\partial q_k} = -\frac{1}{F} \cdot \frac{\partial F_i}{\partial q_k}.$$

Thus the tangential quadric is given by

$$\frac{\partial F_i}{\partial q_k} q_i q_k + F = 0.$$

Substituting for F its expression (17), we obtain for $q_i = 0$

$$F = \left(\int_R \mu dx_1 \cdots dx_n\right)^{-\frac{1}{n+1}}, \qquad \frac{\partial F_i}{\partial q_k} = -(n+2)F^{n+2} \int_R \mu x_i x_k dx_1 \cdots dx_n.$$

Thus the equation of Q in tangential coördinates becomes

$$\int_R \mu dx_1 \cdots dx_n = (n+2) \int_R \mu (x_i q_i)^2 dx_1 \cdots dx_n. \tag{23}$$

Q is by definition covariant under affine transformations. Thus we may assume, without restriction of generality, that Q is the unit sphere $q_1^2 + q_2^2 + \cdots + q_n^2 = 1$. In that case we must have

$$(n+2) \int_R \mu x_i x_k dx_1 \cdots dx_n = \delta_i^k \int_R \mu dx_1 \cdots dx_n. \tag{24}$$

These equations may be interpreted as stating that the moment of inertia of R with respect to any axis is the same as that of the homogeneous unit sphere of same mass as R. Since the property that two regions have the same moment

of inertia with respect to every axis is invariant under affine transformations, we see that irrespective of the equations (24) R has the same moment of inertia as Q with respect to every axis, if Q is covered with homogeneous mass of same total mass as R, i.e., we have proved

5.4. *The tangential quadric at a point P of R is identical with Legendre's ellipsoid of inertia of R, if the polar of mass of P is taken as plane at infinity, i.e., if P is the center of mass of R.*[6]

If we denote as the *generalized* ellipsoid of inertia of R in P a quadric which is covariant under collineations and coincides with Legendre's ellipsoid of inertia in case P is the center of mass of R, the tangential quadric will be identical with the generalized ellipsoid of inertia, i.e., the P.C. with respect to R is in the neighborhood of P approximately given by the P.C. with respect to the generalized ellipsoid of inertia in P.

Let us again make use of the special non-homogeneous coördinate system, in which Q is the unit sphere about P, i.e., in which (24) holds. It then follows that

$$(25) \qquad (n+2)\int_R \mu(x_1^2 + \cdots + x_n^2)dx_1 \cdots dx_n = n\int_R \mu dx_1 \cdots dx_n .$$

From this equation it follows that the inequality

$$x_1^2 + \cdots + x_n^2 < \frac{n}{n+2}$$

cannot hold for all points of R. Hence:

5.5.1. *If R is enlarged in the ratio $(n+2)^{\frac{1}{2}} : n^{\frac{1}{2}}$, it is not contained in Legendre's ellipsoid of inertia of R.*

5.5.1 may be expressed in the projectively invariant form:

5.5.2. *Let Q be the tangential quadric at a point P with polar of mass p. Then there is at least one point S of R, such that the cross ratio $(P, p/S, s) \geqq n/(n+2)$, where s is the polar of S with respect to Q.*

5.5.2 is remarkable, because it gives an inequality for the tangential quadric in a P.C. generated by a mass distribution, which seems not to be satisfied in every positive, symmetric P.C.

6. Definition. The P.C. corresponding to a homogeneous mass-distribution (μ = const.) may be denoted as the *principal* P.C. with respect to R.

6.1.1. *Let P be a point of R and p its polar of mass in the principal P.C. Let*

[6] Legendre's ellipsoid of inertia of a body is defined as the homogeneous ellipsoid having the same moment of inertia with respect to every axis as the body. (Cf. Blaschke, Ber. Verh. sächs. Akad. d. Wiss., vol. 70 (1918), pp. 72–75.) Legendre's ellipsoid is essentially the polar reciprocal of Binet's *fundamental* ellipsoid of inertia, which is characterized by the property that the moment of inertia with respect to any plane through its center is inversely proportional to the square of a radius vector perpendicular to it. (Cf., e.g., A. G. Webster, *Dynamics*, p. 231.)

f be any $(n-2)$-flat contained in p. If then q_f' and q_f'' denote the planes of support of R passing through f and q_f the plane through P and f, then

$$-n \leqq (pq_f/q_f'q_f'') \leqq -\frac{1}{n}.$$

(If R is a quadric, the cross ratio always has the value -1.)

Proof. 6.1.1 is the projective formulation of the following affine theorem given by Minkowski: *the distance of the center of mass of a homogeneous convex body from any plane of support s lies between $B/(n+1)$ and $nB/(n+1)$, where B is the distance of s from the plane of support parallel to s.*[7]

The following dual theorem is proved similarly.

6.1.2. *Given a point P of R and its polar p under the principal P.C. Let F denote any line through P and let Q_F, Q_F', Q_F'' be respectively its points of intersection with p and with the boundary of R. Then*

$$-n \leqq (PQ_F/Q_F'Q_F'') \leqq -\frac{1}{n}.$$

If μ is constant and if a non-homogeneous coördinate system is chosen, in which Legendre's ellipsoid of inertia is the unit sphere about P, according to (25)

$$(n+2)\int_R (x_1^2 + \cdots + x_n^2)dx_1 \cdots dx_n = n\int_R dx_1 \cdots dx_n = nV, \tag{26}$$

V being the volume of R. Let us introduce polar coördinates by $x_i = r\xi_i$, where $(\xi_1, \cdots, \xi_n)$ is a point of the unit sphere Ω. If $d\omega$ is the element of surface of Ω, (26) may be written

$$\int_\Omega \rho^{n+2} d\omega = \int_\Omega \rho^n d\omega, \tag{27}$$

where $r = \rho(\xi_1, \cdots, \xi_n)$ is the equation of the boundary of R. From (27) it follows that ρ can neither be > 1 nor < 1 for all points of the boundary of R. Thus the boundary of R has certainly points in common with the boundary of the unit sphere. This proves the following projectively invariant property of the principal P.C.

6.2. *The boundary of R is intersected by the boundary of every tangential quadric in the principal P.C.*

Compare with 5.5.2 and 4.4.

Some other consequence of (26) may be pointed out. Let again the unit sphere about P be Legendre's ellipsoid of inertia of R. Let Σ be the sphere of volume V about P. Then

$$\int_R r^2 dx_1 \cdots dx_n \geqq \int_\Sigma r^2 dx_1 \cdots dx_n, \tag{28}$$

[7] Cf. Bonnesen and Fenchel, *Konvexe Körper*, p. 52.

as r^2 is less in any point of Σ not in R than in any point of R not in Σ. If ρ_Σ is the radius of Σ, then

$$\int_\Sigma r^2 dx_1 \cdots dx_n = \frac{n}{n+2}\rho_\Sigma^2 \cdot V.$$

Thus according to (26) and (28)

$$V \geqq \rho_\Sigma^2 \cdot V,$$

i.e., $\rho_\Sigma \leqq 1$; this means that the sphere of volume V has no greater radius than the unit sphere, and consequently the volume of the unit sphere is not less than that of R. As Legendre's ellipsoid is covariant under affine transformations, we have thus proved the following theorem of Blaschke:

6.3. *The volume of Legendre's ellipsoid of inertia of R is not less than that of R.*[8]

Every theorem on moments of inertia of homogeneous convex regions may be interpreted as a statement on the tangential quadrics in a principal P.C. Cf., e.g., in this connection the author's paper *Moments of inertia of convex regions*, this Journal, vol. 2, pp. 447–452.

Remark. Given an algebraic surface of degree $2m$ consisting of m closed surfaces containing one another. Let Σ be the (convex) most interior of these surfaces. Then there is a 1–1 correspondence of the points of Σ and the planes outside Σ, if the plane corresponding to a point is its ordinary plane polar with respect to the algebraic surface. This correspondence is a positive, symmetric P.C.

University of Kentucky.

[8] Cf. Blaschke, loc. cit. footnote 6, where this theorem appears as a generalization of Sylvester's four-point-problem. The two volumes in 6.3 can only be equal, if R is an ellipsoid, as the equality sign in (28) holds only if R is a sphere. The convexity of R is not actually used in the proof.

f be any $(n-2)$-flat contained in p. If then q_f' and q_f'' denote the planes of support of R passing through f and q_f the plane through P and f, then

$$-n \leqq (pq_f/q_f'q_f'') \leqq -\frac{1}{n}.$$

(If R is a quadric, the cross ratio always has the value -1.)

Proof. 6.1.1 is the projective formulation of the following affine theorem given by Minkowski: *the distance of the center of mass of a homogeneous convex body from any plane of support s lies between $B/(n+1)$ and $nB/(n+1)$, where B is the distance of s from the plane of support parallel to s.*[7]

The following dual theorem is proved similarly.

6.1.2. *Given a point P of R and its polar p under the principal P.C. Let F denote any line through P and let Q_F, Q_F', Q_F'' be respectively its points of intersection with p and with the boundary of R. Then*

$$-n \leqq (PQ_F/Q_F'Q_F'') \leqq -\frac{1}{n}.$$

If μ is constant and if a non-homogeneous coördinate system is chosen, in which Legendre's ellipsoid of inertia is the unit sphere about P, according to (25)

$$(26) \qquad (n+2)\int_R (x_1^2 + \cdots + x_n^2)dx_1 \cdots dx_n = n\int_R dx_1 \cdots dx_n = nV,$$

V being the volume of R. Let us introduce polar coördinates by $x_i = r\xi_i$, where $(\xi_1, \cdots, \xi_n)$ is a point of the unit sphere Ω. If $d\omega$ is the element of surface of Ω, (26) may be written

$$(27) \qquad \int_\Omega \rho^{n+2} d\omega = \int_\Omega \rho^n d\omega,$$

where $r = \rho(\xi_1, \cdots, \xi_n)$ is the equation of the boundary of R. From (27) it follows that ρ can neither be > 1 nor < 1 for all points of the boundary of R. Thus the boundary of R has certainly points in common with the boundary of the unit sphere. This proves the following projectively invariant property of the principal P.C.

6.2. *The boundary of R is intersected by the boundary of every tangential quadric in the principal P.C.*

Compare with 5.5.2 and 4.4.

Some other consequence of (26) may be pointed out. Let again the unit sphere about P be Legendre's ellipsoid of inertia of R. Let Σ be the sphere of volume V about P. Then

$$(28) \qquad \int_R r^2 dx_1 \cdots dx_n \geqq \int_\Sigma r^2 dx_1 \cdots dx_n,$$

[7] Cf. Bonnesen and Fenchel, *Konvexe Körper*, p. 52.

as r^2 is less in any point of Σ not in R than in any point of R not in Σ. If ρ_Σ is the radius of Σ, then

$$\int_\Sigma r^2 dx_1 \cdots dx_n = \frac{n}{n+2} \rho_\Sigma^2 \cdot V.$$

Thus according to (26) and (28)

$$V \geqq \rho_\Sigma^2 \cdot V,$$

i.e., $\rho_\Sigma \leqq 1$; this means that the sphere of volume V has no greater radius than the unit sphere, and consequently the volume of the unit sphere is not less than that of R. As Legendre's ellipsoid is covariant under affine transformations, we have thus proved the following theorem of Blaschke:

6.3. *The volume of Legendre's ellipsoid of inertia of R is not less than that of R.*[8]

Every theorem on moments of inertia of homogeneous convex regions may be interpreted as a statement on the tangential quadrics in a principal P.C. Cf., e.g., in this connection the author's paper *Moments of inertia of convex regions*, this Journal, vol. 2, pp. 447–452.

Remark. Given an algebraic surface of degree $2m$ consisting of m closed surfaces containing one another. Let Σ be the (convex) most interior of these surfaces. Then there is a 1–1 correspondence of the points of Σ and the planes outside Σ, if the plane corresponding to a point is its ordinary plane polar with respect to the algebraic surface. This correspondence is a positive, symmetric P.C.

UNIVERSITY OF KENTUCKY.

[8] Cf. Blaschke, loc. cit. footnote 6, where this theorem appears as a generalization of Sylvester's four-point-problem. The two volumes in 6.3 can only be equal, if R is an ellipsoid, as the equality sign in (28) holds only if R is a sphere. The convexity of R is not actually used in the proof.

INTERIOR TRANSFORMATIONS ON COMPACT SETS

By G. T. Whyburn

1. As originally defined by Stoilow,[1] a single-valued continuous transformation $T(A) = B$ is called an *interior transformation* provided (1) the image of every open set in A is open in B and (2) the inverse set $T^{-1}(b)$ of each point b in B is totally disconnected. Most later writers have omitted condition (2) and have spoken of an interior transformation as one satisfying (1) alone. While this latter point of view will be adhered to in this paper, it turns out that in most cases the hypotheses in the theorems are such as to make both (1) and (2) satisfied. In order to save words we shall call a continuous transformation satisfying (2) a *light transformation*. Hence a *light interior transformation* in our terminology is the same as an interior transformation as defined by Stoilow.

Our object will be to develop fundamental properties of interior transformations as applied to compact metric sets with a minimum of restriction on the image set. All our sets are assumed to lie in a metric space. We begin with some basic lemmas and theorems.

(1.1) Lemma. *Let $T(R) = S$ be single-valued and let Q be a subset of R such that*

(i) $$Q = T^{-1}T(Q).$$

Then for any subset X of R we have

(ii) $$T(Q\cdot X) = T(Q)\cdot T(X).$$

Proof. Obviously $T(Q\cdot X) \subset T(Q)\cdot T(X)$. To prove the reverse inclusion, let $x \in T(Q)\cdot T(X)$. Then $T^{-1}(x) \subset Q$, $T^{-1}(x)\cdot X \neq 0$. Thus if $y \in T^{-1}(x)\cdot X$, we have $y \in Q\cdot X\cdot T^{-1}(x)$, a relation which gives $T(y) = x \in T(Q\cdot X)$.

(1.2) Lemma. *If $T(R) = S$ is an interior transformation on R, and Q is any subset of R satisfying* (i), *then T is an interior transformation on Q.*

Proof. Let X_q be any open subset of Q. There exists an open subset X of R such that $X_q = X\cdot Q$. By hypothesis $T(X) = U$ is open in S, and by (1.1)

$$T(X_q) = T(Q\cdot X) = T(Q)\cdot T(X) = T(Q)\cdot U.$$

Thus $T(X_q)$ is open in $T(Q)$ and the lemma follows.

(1.3) Lemma. *If A is compact and $T(A) = B$ is continuous, then for any set $R \subset A$*

(i) $$\overline{T(R)} = T(\bar{R}),$$

(ii) $$\overline{T(R)} - T(R) \subset T(\bar{R} - R).$$

Received March 19, 1937.

[1] See Stoilow, Annales Scientifiques de l'Ecole Normale Supérieure, vol. 63 (1928), pp. 347–382 and Annales de l'Institut Henri Poincaré, vol. 2 (1932), pp. 233–266.

Thus if T is an interior transformation and R is open in A,

(iii) $$F[T(R)] \subset T[F(R)],$$

where in general $F(X)$ denotes the boundary of X.

(1.31) COROLLARY. *The Menger-Urysohn order of a point in A is never increased when A undergoes an interior transformation. The properties of being* (i) *a curve of order $\leqq n$,* (ii) *a regular curve,* (iii) *a rational curve, are invariant under interior transformations.*

(1.4) THEOREM. *Let A and B be compact and let $T(A) = B$ be interior. Then if R is any quasi-connected open set in B, every quasi-component of Q of $T^{-1}(R)$ maps onto all of R under T.*

Proof. Suppose, on the contrary, that for some $p \epsilon R$, $T^{-1}(p) \cdot Q = 0$. Then since $T^{-1}(p)$ is compact and is $\subset T^{-1}(R)$, it follows by an application of the Borel theorem that there exists an open set U such that

(i) $$Q \subset U \subset T^{-1}(R), \qquad F(U) \cdot T^{-1}(R) = 0, \qquad U \cdot T^{-1}(p) = 0.$$

Now $T(U) = V$ is open, and we have

(ii) $$V \supset T(Q) \subset R, \qquad V \subset R - p,$$

whence, since R is quasi-connected, $F(V) \cdot R \neq 0$. But by Lemma (1.3), $F(V) \subset T[F(U)]$ and by (i), $T[F(U)] \cdot R = 0$. Thus our theorem is proved.

(1.41) COROLLARY. *If A is locally connected, if B_0 is any closed set in B and R is any component of $B - B_0$, then $T^{-1}(R)$ has just a finite number of components and each one of these components maps onto all of R under T.*

(1.42) COROLLARY. *If A is locally connected and B is connected and cyclic (i.e., without cut points), then for each $x \epsilon B$, $A - T^{-1}(x)$ has just a finite number of components and each of these components maps onto $B - x$ under T.*

(1.5) THEOREM. *Let $T(A) = B$ be an interior transformation, where A is compact, and let C be any continuum in B. Then for each component K of $T^{-1}(C) = Q$, we have $T(K) = C$.*

For by (1.2), the transformation $T(Q) = C$ is interior. Thus, since C is connected and open in C, it follows by (1.4) that each component of Q maps onto all of C under T.

(1.6) *Let $T(A) = T_2T_1(A) = B$, where $T_1(A) = A'$ and $T_2(A') = B$ are continuous. If T is interior, so also is T_2.*

For let R be any open set in A' and let $Q = T_1^{-1}(R)$. Then since T_1 is continuous, Q is open in A; and since T is interior, $T(Q)$ is open in B. But $T(Q) = T_2T_1(Q) = T_2(R)$. Accordingly, $T_2(R)$ is open in B.

(1.61) COROLLARY. *If A is compact, any interior transformation $T(A) = B$ can be factored into the form $T(A) = T_2T_1(A)$, where T_1 is monotone*[2] *and T_2 is interior and light.*

[2] A transformation $T(A) = B$ is *monotone* provided that for each $b \epsilon B$, $T^{-1}(b)$ is connected.

For T can be factored into the form T_2T_1, where T_1 is monotone and T_2 is light by a known theorem;[3] and by (1.6), T_2 is necessarily interior.

2. **Mapping of the Betti groups.** Eilenberg[4] has raised the question whether under an interior transformation $T(X) = Y$, where X is compact, the Betti groups of X map homomorphically onto the corresponding groups of Y. With respect to the integral coefficient domain $\mathfrak{G}$, this question is easily answered in the negative by the transformation $w = z^2$ of the circle $|z| = 1$ into the circle $|w| = 1$, since in this case clearly[5] $B^1_{\mathfrak{G}}(X)$ maps onto $2B^1_{\mathfrak{G}}(Y)$.

However, if we consider the Betti groups relative to the rational coefficient field R, this question is not nearly so easy to answer. In the above case, clearly $B^1_R(X)$ maps homomorphically onto $B^1_R(Y)$, and indeed it will be proved in the next section that this is always true for any linear graph X undergoing an interior transformation.[6]

The following example is of interest in this connection as well as in connection with a theorem to be proved below in §4.

(2.1) Example. *There exists a compact continuum K with a vanishing one-dimensional integral Betti group and an interior transformation $f(K) = J$ of K onto a circle J.*

To exhibit this, we employ a continuum K which has been constructed by Vietoris[7] for a different purpose. This continuum is constructed by first taking a Cantor ternary set on both the upper and lower bases of the unit square with vertices (0, 0), (1, 0), (1, 1), (0, 1) and joining corresponding points in these two sets (i.e., points with the same abscissa) by the altitude of the square they determine. This gives a set C which may be considered a "Cantor set of unit intervals". K is then constructed from C as follows. We identify the lower endpoints of the left half of these intervals with the upper endpoints of the right half in the same sense from left to right; then divide the remaining lower endpoints in half and identify the left half with the upper right half of the remaining upper endpoints in the same sense; next divide the remaining lower endpoints in half and identify the lower left half with the upper right half of the remaining upper endpoints, and so on indefinitely. Finally identify the points (1, 0) and (0, 1).

Vietoris (loc. cit.) has shown that the one-dimensional integral Betti group of this continuum K reduces to the null element.

We now proceed to define an interior transformation f which will map K

[3] See Eilenberg, Fundamenta Mathematicae, vol. 22 (1934), p. 292 and G. T. Whyburn, American Journal of Mathematics, vol. 56 (1934), p. 297.

[4] See Fundamenta Mathematicae, vol. 24 (1935), p. 175.

[5] We employ here the notation (see Alexandroff-Hopf, *Topologie*) $B^r_J(K)$ for the r-dimensional Betti group of a complex (or set) K relative to a coefficient domain J.

[6] Results obtained recently indicate that this also holds for any compact set X with $p^1(X)$ finite. This is contrary to a statement made by the author in an abstract in the Bulletin of the American Mathematical Society, vol. 43 (1937), p. 183.

[7] See Mathematische Annalen, vol. 97 (1927), p. 459.

onto the unit circle J. Let $f_1(C) = K$ be the (continuous) transformation representing the above "identifying" construction of K from the set C. Let $p \in K$ and let $(x, y) \in f_1^{-1}(p)$, where x and y are Cartesian coördinates. We define

$$f(p) = (x', y'),$$

where

$$\begin{matrix} x' = \cos \pi y \\ y' = \sin \pi y \end{matrix} \quad \text{for } x < \tfrac{1}{2}, \text{ and } \quad \begin{matrix} x' = \cos \pi(1 + y) \\ y' = \sin \pi(1 + y) \end{matrix} \quad \text{for } x > \tfrac{1}{2}.$$

Clearly this transformation is interior and light and it maps K onto the circle $x'^2 + y'^2 = 1$.

3. Linear graphs.

(3.1) THEOREM. *Let A be a linear graph and let $f(A) = B$ be interior. Then B is a linear graph and there exist subdivisions of A and B respectively into finite complexes K_a and K_b such that f maps each edge of K_a topologically into an edge in K_b. Thus $f(K_a) = K_b$ is a simplicial transformation. Furthermore, if J is any simple closed curve in B, there exists a simple closed curve C in A such that $f(C) = J$ and, on C, f is topologically equivalent[8] to the transformation $w = z^k$ on $|z| = 1$ (k an integer). Consequently $B_R^1(K_a)$ maps homomorphically onto $B_R^1(K_b)$ under f.*

Proof. Since A has only a finite number of points of order > 2 and has no point of increasing order, and since, by (1.31), the order of no point can be increased under f, it follows that B has these same properties. Accordingly, B is a linear graph.

Furthermore, for each $b \in B$, $f^{-1}(b)$ must be a finite set of points. For otherwise there would exist infinitely many components of $A - f^{-1}(b)$, and clearly this is impossible, since by (1.41) the inverse of each one of the finite number of components of $B - b$ is a finite set of components of $A - T^{-1}(b)$.

Now let E_a and E_b respectively denote the sets of points in A and B of order $\neq 2$. Let $D_b = E_b + f(E_a) +$ one point from each component of $B - [E_b + f(E_a)]$ which has at most one limit point in $E_b + f(E_a)$; let $D_a = f^{-1}(D_b)$. The sets D_a and D_b subdivide A and B into complexes K_a and K_b, respectively. Let xy be any edge of K_a and let $f(x) = x'$, $f(y) = y'$, $f(xy) = x'y'$. Then $x' \neq y'$, since otherwise $f(xy - x - y)$ would be a free arc in B having just one endpoint, which is impossible by the choice of D_b. Furthermore, for any $p \in x'y' - x' - y'$, $f^{-1}(p) \cdot xy$ reduces to just one point. For if not, there would exist an open arc $\widehat{p_1p_2}$ in $xy - xy \cdot f^{-1}(p)$ with $p_1 + p_2 \subset f^{-1}(p)$; this would give $f(\widehat{p_1p_2}) \supset x'$ or $f(\widehat{p_1p_2}) \supset y'$; and both are impossible, since $f^{-1}(x') \cdot xy = x$,

[8] Two transformations $T(A) = B$ and $W(A') = B'$ are said to be topologically equivalent provided there exist topological transformations $H_1(A) = A'$ and $H_2(B') = B$ such that $T(x) = H_2WH_1(x)$ for every $x \in A$. See my paper *Completely alternating transformations*, Fundamenta Mathematicae, vol. 27 (1936), pp. 140–146.

$f^{-1}(y')\cdot xy = y'$. Accordingly, $f(xy) = x'y'$ is topological and $f(K_a) = K_b$ is simplicial.

Now let J be any simple closed curve in B, let the vertices of K_b on J be ordered cyclically $p_0, p_1, \cdots, p_n, p_0$ and let $q_0 \epsilon f^{-1}(p_0)$. There exists an edge q_0q_1 of K_a which is contained in $f^{-1}(p_0p_1)$ and such that $q_1 \epsilon f^{-1}(p_1)$. Similarly there is an edge q_1q_2 with $q_1q_2 \subset f^{-1}(p_1p_2)$ and $q_2 \epsilon f^{-1}(p_2)$, and so on to p_n. There is an edge $q_nq_0^1$ with $q_nq_0^1 \subset f^{-1}(p_np_0)$ and $q_0^1 \epsilon f^{-1}(p_0)$. If $q_0^1 = q_0$, we can bring our selection to a close. If not, we choose an edge $q_0^1q_1^1$ in $f^{-1}(p_0p_1)$. Again if $q_1^1 = q_1$, we may stop. Otherwise we choose an edge $q_1^1q_2^1$ in $f^{-1}(p_1p_2)$, and so on. Continuing this process, since $f^{-1}(b)$ is a finite set for each $b \epsilon B$, after a finite number of steps we must eventually reach, for the first time, a point q_i^j $(i \leqq n)$ such that $q_i^j = q_i^m$ $(m < j)$. Then clearly the edges

$$q_0^{m+1}q_1^{m+1},\ q_1^{m+1}q_2^{m+1},\ \cdots,\ q_{n-1}^{m+1}q_n^{m+1},\ q_n^{m+1}q_0^{m+2},\ \cdots,\ q_{i-1}^jq_i^j,\ q_i^jq_{i+1}^m,$$

$$q_{i+1}^mq_{i+2}^m,\ \cdots,\ q_n^mq_0^{m+1}$$

fit together to form a simple closed curve C which maps onto J exactly $j - m$ times; and since $q_s^rq_{s+1}^r$ maps topologically onto p_sp_{s+1}, it follows at once that if we set $j - m = k$, the transformation $f(C) = J$ is topologically equivalent to the transformation $w = z^k$ on the circle $|z| = 1$ of the complex z-plane.

Since, as shown above, f can be considered as a simplicial transformation of K_a onto K_b, it follows that f generates a homomorphism, which we shall also call f, of $B_R^1(K_a)$ into $B_R^1(K_b)$. To see that actually $f[B_R^1(K_a)] = B_R^1(K_b)$, let us take any rational 1-cycle z in K_b. As is well known, z can be expressed as a linear form

$$z = a_1z_1 + a_2z_2 + \cdots + a_nz_n \qquad (a_i \text{ rational}),$$

where z_i is a simply oriented simple closed polygon. By the above, we can find a simple closed curve C_i in A such that $f(c_i) = k_iz_i$ $(i \leqq n,\ k_i$ an integer), where c_i is an oriented fundamental cycle on C_i. Accordingly if we set

$$\gamma = \frac{a_1}{k_1}c_1 + \frac{a_2}{k_2}c_2 + \cdots + \frac{a_n}{k_n}c_n,$$

then γ is a rational 1-cycle in K_a and $f(\gamma) = z$. Thus the homology class of γ in $B_R^1(K_a)$ maps onto the homology class of z in $B_R^1(K_b)$ and our result follows.

(3.11) COROLLARY. $p^1(K_b) \leqq p^1(K_a)$ (*where* $p^r(X)$ *denotes the r-dimensional Betti number of* X).

(3.2) *If A is a simple closed curve and $T(A) = B$ is interior, then B is either a simple closed curve or a simple arc. If B is a simple closed curve, there exists an integer k such that T is topologically equivalent to the transformation $w = z^k$ on the circle $|z| = 1$. If B is an arc, there exists an integer k such that T is equivalent to the transformation $f(1, \theta) = \sin k\theta/2$ of the circle $\rho = 1$ into the interval* $(-1, 1)$.

Proof. It follows from (1.31) that every point of B is of order $\leqq 2$. Accordingly, B is either a simple closed curve or a simple arc.

If B is a simple closed curve, it follows from (3.1) by taking $J = B$ that T is equivalent to $w = z^k$ on $|z| = 1$ for some integer k.

If B is a simple arc xy, referring back to the proof of (3.1), we see that $D_b = x + y$ and hence $D_a = T^{-1}(x) + T^{-1}(y)$. Thus if we order the points of D_a cyclically on A: $p_1, p_2, \cdots, p_k, p_1$, where $p_1 \epsilon T^{-1}(x)$, then by (3.1), p_1p_2 maps topologically onto xy, p_2p_3 maps topologically onto yx, p_3p_4 onto xy and so on to p_kp_1, which maps onto yx. (Note that k must be even.) Hence T is equivalent to the transformation $f(1, \theta) = \sin \frac{1}{2}k\theta$ of $\rho = 1$ into $(-1, 1)$.

Remark. In each case the integer k is exactly the maximum power among the sets $T^{-1}(b)$ for all $b \epsilon B$. If k is odd, B is necessarily a simple closed curve, T is completely alternating[9] and all the sets $T^{-1}(b)$ are of power k. If k is even, B may be either a simple closed curve or an arc xy; and in the latter case, $T^{-1}(x)$ and $T^{-1}(y)$ are each of power $\frac{1}{2}k$, while for every other point b of B, $T^{-1}(b)$ is of power k.

(3.3) *If A is a simple arc pq and $T(A) = B$ is interior, B is a simple arc and there exists an integer k such that T is topologically equivalent to the transformation $f(x) = \cos k\pi x$ of the interval* $(0, 1)$ *into the interval* $(-1, 1)$.

Proof. Since, by (1.1), B must have at least two points of order 1 but no point of order > 2, it follows that B is a simple arc uv.

Referring again to the proof of (3.1), we see that in this case $D_b = u + v \supset T(p) + T(q)$ and $D_a = T^{-1}(u) + T^{-1}(v)$. Hence if we order the points of D_a on A from p to q: $p = p_0, p_1, \cdots, p_k = q$ and suppose the notation chosen so that $T(p) = u$, then by (3.1), p_0p_1 maps topologically onto uv, p_1p_2 maps topologically onto vu, p_2p_3 onto uv, and so on to $p_{k-1}p_k$, which may map either onto uv or vu. Hence it is seen at once that T is equivalent to the transformation $f(x) = \cos k\pi x$ of $(0, 1)$ into $(-1, 1)$.

By Corollary (3.11), the Betti numbers of a graph cannot be increased under an interior transformation. This conclusion can be strengthened somewhat for graphs without endpoints, namely, in that we can obtain the same conclusion for any continuous transformation $T(A) = B$ which increases the (Menger-Urysohn) order of no point of A. Such a transformation will be called *non-order-increasing*.

(3.4) Lemma. *If A is a connected linear graph without endpoints and $T(A) = B$ is a non-order-increasing transformation, then B is a linear graph and $p^1(B) \leq p^1(A)$.*

Proof. Since A has only a finite number of points of order $\neq 2$ and has no point of increasing order, it follows that B must have these same properties and hence must be a graph.

Let X be a finite set of points in A such that X contains all points of A of order $\neq 2$ and $T(X) = Y$ contains all points of B of order $\neq 2$ and such that each component of $A - X$ or $B - Y$ has two distinct limit points in X or Y respectively. Let α^0 and β^0 be the number of points in X and Y respectively. Let α^1 and β^1 be the number of components in $A - X$ and $B - Y$ respectively.

[9] Loc. cit.

Then by the Euler-Poincaré formula we have

$$\text{(i)} \qquad p^1(A) = \alpha^1 - \alpha^0 + 1, \qquad p^1(B) = \beta^1 - \beta^0 + 1.$$

Also, if we set

$$X = \sum_1^{\alpha^0} x_i, \qquad Y = \sum_1^{\beta^0} y_i,$$

we have

$$\text{(ii)} \qquad \sum_1^{\alpha^0} o(x_i) = 2\alpha^1, \qquad \sum_1^{\beta^0} o(y_i) = 2\beta^1,$$

where in general $o(x)$ denotes the Menger-Urysohn order of the point x. Now for each $i \leqq \beta^0$, let

$$X \cdot T^{-1}(y_i) = x_i^1 + x_i^2 + \cdots + x_i^{k_i}.$$

Then

$$\text{(iii)} \qquad \sum_1^{\beta^0} k_i = \alpha^0 \quad \text{and} \quad \sum_{i=1}^{\beta^0} \sum_{j=1}^{k_i} x_i^j = X = \sum_1^{\alpha^0} x_i.$$

Since by hypothesis $o(x_i^j) \geqq o(y_i)$ and $o(x_i^j) \geqq 2$, we have

$$\text{(iv)} \qquad \sum_{j=1}^{k_i} o(x_i^j) \geqq o(y_i) + \sum_{j=2}^{k_i} o(x_i^j) \geqq o(y_i) + 2(k_i - 1),$$

whence

$$\text{(v)} \qquad \sum_{i=1}^{\beta^0} \sum_{j=1}^{k_i} o(x_i^j) \geqq \sum_{i=1}^{\beta^0} o(y_i) + 2 \sum_{i=1}^{\beta^0} k_i - 2\beta^0.$$

Thus from (ii), (iii) and (v) we get

$$2\alpha^1 \geqq 2\beta^1 + 2\alpha^0 - 2\beta^0$$

or

$$(\alpha^1 - \beta^1) - (\alpha^0 - \beta^0) \geqq 0,$$

whence, from (i),

$$p^1(A) - p^1(B) = (\alpha^1 - \beta^1) - (\alpha^0 - \beta^0) \geqq 0$$

or

$$p^1(A) \geqq p^1(B).$$

(3.5) *Let $T(A) = B$ be non-order-increasing. Let A be a connected graph having a set X of q endpoints and let $T(X)$ contain r points. Then B is a graph and $p^1(B) \leqq p^1(A) + (\frac{1}{2})(q - r)$.*

For we can replace (iv) by

$$\text{(iv)}' \qquad \sum_{j=1}^{k_i} o(x_i^j) \geqq o(y_i) + \sum_{j=2}^{k_i} o(x_i^j) \geqq o(y_i) + 2(k_i - 1) - (q_i - 1) \quad (i \leqq r),$$

where q_i is the number of points of order 1 of A in $T^{-1}(y_i)$ and the notation is chosen so that $o(x_i^1) = \min [o(x_i^j)]$ and $q_i > 0$ for $i \leqq r$. Summing, we get

$$\text{(v)}' \qquad \sum_{i=1}^{\beta^0} \sum_{j=1}^{k_i} o(x_i^j) = \sum_{i=1}^{\beta^0} o(y_i) + 2 \sum_{i=1}^{\beta^0} k_i - 2\beta^0 - \sum_{i=1}^{\gamma} q_i + r.$$

Whence, as before,

$$2\alpha^1 \geqq 2\beta^1 + 2\alpha^0 - 2\beta^0 - (q - r),$$
$$(\alpha^1 - \beta^1) - (\alpha^0 - \beta^0) \geqq -\tfrac{1}{2}(q - r),$$
$$p^1(A) - p^1(B) \geqq -\tfrac{1}{2}(q - r),$$
$$p^1(B) \leqq p^1(A) + \tfrac{1}{2}(q - r).$$

(3.51) Corollary. *If*
(i) $q = 0$, *i.e., there are no endpoints,*
(ii) $q = r$, *i.e., T is (1-1) on the set X of endpoints of A,*
(iii) $q - r \leqq 1$, *i.e., at most one pair of endpoints in A map into one point in B, or*
(iv) $q \leqq 2$, *i.e., there are at most two endpoints in A,*
then $p^1(B) \leqq p^1(A)$.

(3.6) Theorem. *If A is a one-dimensional compact locally connected continuum and $T(A) = B$ is an interior transformation, then $p^1(B) \leqq p^1(A)$.*

Proof. Clearly we may suppose $p^1(A)$ finite. If, contrary to our theorem, $p^1(B) > p^1(A)$, we can choose an A-set H in B without endpoints and such that $p^1(H) > p^1(A)$, for we need only take a finite number of true cyclic elements of B so that their sum S has Betti number $p^1(S) > p^1(A)$ and let H be the smallest A-set in B containing S.

Let K be a component of $T^{-1}(H)$. Then since $T(K) = H$ is an interior transformation [by (1.5) above, $T(K) = H$; and T is interior on K, since there are only a finite number of components of $T^{-1}(H)$], and since H has no endpoints, it follows that K has no endpoints because T is non-order-increasing on K. Whence, by (3.4), $p^1(H) \leqq p^1(K)$. But $p^1(K) \leqq p^1(A)$, and this gives $p^1(H) \leqq p^1(K) \leqq p^1(A)$, contrary to our supposition. Accordingly $p^1(B) \leqq p^1(A)$.

(3.61) Corollary. *If A is a dendrite, so also is B.*

4. **Light interior transformations.** It will be recalled from the introduction that a transformation $T(A) = B$ is *light* provided that for each $b \in B$, $T^{-1}(b)$ is totally disconnected. We begin this section with a theorem for arbitrary compact sets analogous to a result of Stoilow's[10] for the case of a transformation of one plane region into another.

(4.1) Theorem. *Let $T(A) = B$ be a light interior transformation, where A is compact. Then if pq is any simple arc in B and p_0 is any point in $T^{-1}(p)$, there exists a simple arc p_0q_0 in A such that $T(p_0q_0) = pq$ and T is topological on p_0q_0.*

[10] See Stoilow, loc. cit. The author learned recently that a theorem essentially the same as (4.1) has been discovered independently by Zippin and Montgomery. See an article by Montgomery in a forthcoming issue of the Transactions of the American Mathematical Society.

Our proof for this theorem will make use of the following definition and lemma.

DEFINITION. If $\sigma:p = x_0, x_1, \cdots, x_n = q$ (x_i precedes x_{i+1} in the order p, q) is a subdivision of pq, a continuum K in $T^{-1}(pq)$ will be said to proceed directly from $T^{-1}(p)$ to $T^{-1}(q)$ relative to the subdivision σ provided $T(K) = pq$ and $K = \sum_1^n K_i$, where K_i is a continuum in $T^{-1}(x_{i-1}x_i)$. By the *norm* of such a continuum K will be meant max $\delta(K_i)$.

LEMMA. *Given any subdivision σ' of pq and any $\epsilon > 0$, there exists a subdivision σ of pq containing σ' and a continuum K of norm $< \epsilon$ in $T^{-1}(pq)$ which contains p_0 and which proceeds directly from $T^{-1}(p)$ to $T^{-1}(q)$ relative to σ.*

Proof of Lemma. Let S denote the set of all points q' on pq for which such a subdivision exists on pq' containing all points of σ' on pq'. Let x be any point of pq.

If $x \epsilon S$, let us take the subdivision $\sigma = \sigma_x$ on px and the corresponding continuum $K = K_x$. Let $z \epsilon K \cdot T^{-1}(x)$. There exists an ϵ-neighborhood U of z such that $F(U) \cdot T^{-1}(x) = 0$. There exists a point v on x_q such that if w is any point whatever of xv, $F(U) \cdot T^{-1}(w) = 0$ and w does not belong to σ' unless $w = x$. It follows that $Q = T^{-1}(xv) \cdot U$ is both open and closed in $T^{-1}(xv)$. Accordingly $T(Q) = xv$ is an interior transformation. Let K_{n+1} be the component of Q containing z, let σ_v be the subdivision $p = x_0, x_1, \cdots, x_n, x_{n+1} = v$ of pv and let $K_v = \sum_1^{n+1} K_i$. Then clearly K_v is a continuum of norm $< \epsilon$ which proceeds directly from $T^{-1}(p)$ to $T^{-1}(v)$ relative to σ_v, so that v belongs to S. Thus S is open in pq.

Now if $x \epsilon pq - S$, we can choose x to be the first point of $pq - S$ on pq. Hence $px - x \subset S$. If we take a sequence $y_1, y_2, \cdots$ on px so that $y_i \rightarrow x$, and choose points $z_i \epsilon K_{y_i} \cdot T^{-1}(y_i)$, where K_{y_i} is a continuum of norm $< \epsilon$ in $T^{-1}(py_i)$ proceeding directly from $T^{-1}(p)$ to $T^{-1}(y_i)$ relative to a subdivision σ_{y_i} of py_i which contains all points of σ' on py_i, the sequence $z_1, z_2, \cdots$ will have a limit point z on $T^{-1}(x)$, and we may suppose $z_i \rightarrow z$. Let us choose the ϵ-neighborhood U of x exactly as before. There exists an i such that for any $w \epsilon y_i x$, $T^{-1}(w) \cdot F(U) = 0$ and such that no point of $y_i x$, except possibly x, belongs to σ'. Hence if $Q = T^{-1}(y_i x) \cdot U$, Q is both open and closed in $T^{-1}(y_i x)$ and $T(Q) = y_i x$ is interior. Now if σ_{y_i} is denoted by $p = x_0, x_1, \cdots, x_n = y_i$ and $K_{y_i} = \sum_1^n K_m$, let σ_x be the subdivision $p = x_0, x_1, \cdots, x_n, x_{n+1} = x$ of px, let K_{n+1} be the component of Q containing z_i and $K_x = \sum_1^{n+1} K_j$. Then clearly K_x is a continuum of norm $< \epsilon$ which proceeds directly from $T^{-1}(p)$ to $T^{-1}(x)$ relative to σ_x, and σ_x contains all points of σ' on px. But this gives $x \epsilon S$, contrary to our supposition. Accordingly $S = pq$, and the lemma is proved.

Proof of the Theorem. There exists a subdivision σ_1: $p = x_0^1, \cdots, x_{n_1}^1 = q$ of pq and a continuum $K^1 = \sum_1^{n_1} K_i^1$ of norm < 1 which contains p_0 and proceeds directly from $T^{-1}(p)$ to $T^{-1}(q)$ relative to σ_1.

In general, for each k there exists a subdivision

$$\sigma_k: p = x_0^k, \cdots, x_{n_k}^k = q,$$

which contains σ_{k-1}, and a continuum $K^k = \sum_1^{n_k} K_i^k$ of norm $< 1/k$ which contains p_0 and proceeds directly from $T^{-1}(p)$ to $T^{-1}(q)$ relative to σ_k.

Since the sequence $K^1, K^2, \cdots$ contains a convergent subsequence, we may suppose, without loss of generality, that the whole sequence converges to a limit which we shall call E. Since each K^i is a continuum with $T(K^i) = pq$, it follows that E is a continuum and $T(E) = pq$. To prove that E is an arc and $T(E) = pq$ is topological, it suffices to prove that for each $y \in pq$, $T^{-1}(y) \cdot E$ reduces to a single point.

Now since for each k, i, j ($i < j, i, j \leqq n_k$) there exists a continuum $K_{ij}^k = \sum_{m=i+1}^{j} K_m^k$ in K^k which contains $K^k \cdot T^{-1}(\widehat{x_i^k x_j^k})$, it follows that for each k, i, j there exists a continuum $E_{ij}^k = \lim_{(m)} K_{i_m j_m}^{k+m}$ in E which contains $E \cdot T^{-1}(\widehat{x_i^k x_j^k})$. (Note: i_m and j_m are integers such that $x_{i_m}^{k+m} = x_i^k$, $x_{j_m}^{k+m} = x_j^k$.)

Now for any $y \in pq$ and any k let us choose i and j so that (i) y is interior to the arc $x_i^k x_j^k$ if $p \neq y \neq q$, (ii) $y = x_i^k = x_0^k = p$ if $y = p$, $y = x_j^k = q$ if $y = q$, and (iii) $0 < j - i \leqq 2$. Then since in all cases $E \cdot T^{-1}(y)$ is interior to E_{ij}^k (rel. E), it follows that $E \cdot T^{-1}(y) = \lim_{(k \to \infty)} E_{ij}^k$. Accordingly, since each E_{ij}^k is a continuum, $E \cdot T^{-1}(y)$ is a continuum. Therefore $E \cdot T^{-1}(y)$ reduces to a single point, since $T^{-1}(y)$ is totally disconnected, and our theorem is proved.

Remarks. It is of interest to note that the hypothesis of the theorem just proved is satisfied even in cases such as that exhibited in the example given in §2. Also it results at once from this theorem that the property of containing no arc is invariant for compact sets under light interior transformations. It may be noted further that, whereas the initial point p_0 of the arc $p_0 q_0$ can be chosen arbitrarily in $T^{-1}(p)$, the terminal point q_0 is not subject to choice but may easily depend on the choice of p_0.

The methods of proof used in (3.1) together with the theorem just proved yield at once

(4.2) THEOREM. *Under the conditions of* (4.1), *if J is any simple closed curve in B, then for any integer $n > 0$ either there exists a simple closed curve C in A and an integer $k \leqq n$ such that $T(C) = J$ and, on C, T is topologically equivalent to the transformation $w = z^k$ on $|z| = 1$, or there exists a simple arc X in A such that $T(X) = J$ and, on X, T is equivalent to the transformation $w = e^{yi}$ on the interval $-(n+1)\pi \leqq y \leqq (n+1)\pi$. If there exists no such simple closed curve C for any k, then there exists an open curve L in A such that $T(L) = J$ and, on L, T is equivalent to the transformation $w = e^{yi}$ on $-\infty < y < \infty$.*

Note. Since on a linear graph A any interior transformation is necessarily light, it is clear that in so far as it concerns simple closed curves in B, (3.1) is a consequence of (4.2). Also since the open curve L (when it exists) necessarily

contains infinitely many disjoint intervals all of diameter greater than some $d > 0$, we have

(4.21) COROLLARY. *If $T(A) = B$ is interior and light, where A is hereditarily locally connected, then for any simple closed curve J in B there exists a simple closed curve C in A and an integer k such that $T(C) = J$ and, on C, T is topologically equivalent to the transformation $w = z^k$ on $|z| = 1$.*

(4.3) THEOREM. *Let $T(A) = B$ be interior, where A is compact. For any simple arc pq in A there exists a continuum E in A such that $T(E) = pq$ and, on E, T is monotone.*

Proof. Let us factor T into the form T_2T_1 where T_1 is monotone and T_2 is light. Let $A' = T_1(A)$. Then, by (1.6), $T_2(A') = B$ is interior and light. Hence by (4.1) there exists an arc p_0q_0 in A' such that $T_2(p_0q_0) = pq$ is topological. Let $E = T_1^{-1}(p_0q_0)$. Then E is a continuum since T_1 is monotone, and since

$$T(E) = T_2T_1(E) = T_2T_1T_1^{-1}(p_0q_0) = T_2(p_0q_0) = pq$$

and T_2 is topological on $T_1(E) = p_0q_0$, it follows that $T(E) = pq$ is monotone.

Note. Theorem (4.1) also follows directly from this theorem.

(4.4) THEOREM. *Let $T(A) = B$ be interior and light, where A is compact. If K is any locally connected continuum in B and C is any component of $T^{-1}(K)$, the transformation $T(C) = K$ is interior.*

Proof. Let E be any open subset of C. There exists an open subset D of $T^{-1}(K)$ such that $D \cdot C = E$. Let y be any point of $T(E)$ and let $x \in E \cdot T^{-1}(y)$. Since dim $T^{-1}(y) = 0$, there exists a set U open in D which contains x and is contained in D and such that $F(U) \cdot T^{-1}(y) = 0$. Let $T[F(U)] = K_0$, let R be the component of $K - K_0$ containing y and let Q be the quasi-component of $T^{-1}(R)$ containing x. By (1.4), $T(Q) = R$; and since $T^{-1}(R) \subset T^{-1}(K) - T^{-1}(K_0) \subset T^{-1}(K) - F(U)$ and $Q \supset x \in U$, we have $Q \subset U \subset D$. Whence, $Q \subset D \cdot T^{-1}(K)$; and since Q is quasi-connected, this gives $Q \subset C$, so that $Q \subset D \cdot C = E$. Accordingly $T(E) \supset T(Q) \supset R$, so that $T(E)$ is open in K.

5. **Transformations on plane sets.** For light interior transformations applied to plane locally connected continua we prove the following result concerning the invariance of local connectivity under the inverse transformation.

(5.1) THEOREM. *Let A be a plane, compact, locally connected continuum and let $T(A) = B$ be a light interior transformation. Then if K is a locally connected subcontinuum of B, $T^{-1}(K)$ is locally connected.*

Proof. If this is not so, there exists[11] a point $p \in T^{-1}(K) = H$, a simple closed curve J about p with interior E and an infinite sequence $M_1, M_2, \cdots$ of quasi-components of $E \cdot H$ converging to a continuum M in $J + E$ containing p and such that $M_1, M_2, \cdots$ lie in distinct components $N_1, N_2, \cdots$ of $H \cdot (E + J)$

[11] For details of how to obtain these sets the reader is referred to papers by R. L. Moore in the Bulletin of the American Mathematical Society, vol. 25 (1919), pp. 174–176, and vol. 29 (1923), pp. 289–302.

and such that for each i, $k > 0$, $\bar{M}_i$ separates $\bar{M}_{i-1}$ and $\bar{M}_{i+k}$ in $J + E$. Since by hypothesis $T^{-1}T(p)$ is totally disconnected, we may suppose J chosen so that $J \cdot T^{-1}T(p) = 0$. Let $B_0 = T(A \cdot J)$ and let R be the component of $K - K \cdot B_0$ containing $T(p)$. [Note that $T(p)$ cannot belong to B_0, since $J \cdot T^{-1}T(p) = 0$.]

Now by the local connectedness of K, R is open in K. Hence for almost all i, $T(M_i) \cdot R \neq 0$, and thus we may suppose this holds for all i. Since $T(H) = K$ is interior [by (1.2)], it follows by (1.4) that each quasi-component of $T^{-1}(R)$ maps onto all of R under T. Since $T^{-1}(R) \cdot J = 0$, it follows at once that there exists a sequence $L_1, L_2, \cdots$ of distinct quasi-components of $T^{-1}(R)$ converging to a continuum L where $p \in L \subset M$ and for each i, $L_i \subset M_i$.

Since A is locally connected at p but H is not locally connected at p, it follows that there exists a component Q of $B - (K + B_0)$ which has a limit point q in R. Since $T(L_i) = R$ for each i, it follows that for each i there exists a point $q_i \in T^{-1}(q) \cdot L_i$. Now since, by (1.41), $T^{-1}(Q)$ consists of a finite number of components of $A - H - T^{-1}(B_0)$ and each q_i is a limit point of $T^{-1}(Q)$, it follows that there exists some component W of $T^{-1}(Q)$ such that at least three, say $q_{i_1}, q_{i_2}, q_{i_3}$ $(i_1 < i_2 < i_3)$ of the points q_i are limit points of W. Since $T^{-1}(B_0) = A \cdot J$ and $q_i \subset E$, we have $W \subset E$, and since $H \supset M_i$ for each i, we have $W \cdot M_i = 0$ for each i. But now $\bar{M}_{i_2}$separates $\bar{M}_{i_1}$ and $\bar{M}_{i_3}$ in $J + E$, and since $W \subset E$, $\bar{W} \supset q_{i_1} + q_{i_3}$, this gives $W \cdot M_{i_2} \neq 0$, which is impossible. Thus the supposition that our theorem is false leads to a contradiction.

(5.11) Corollary. *Under the conditions of* (5.1) *there are only a finite number of components* $C_1, C_2, \cdots, C_n$ *of* $T^{-1}(K)$ *and, for each* i, $T(C_i) = K$ *is interior.*

It may be shown by simple examples that (5.1) does not remain valid if we remove either the condition that A lie in a plane or that T be a light transformation.

University of Virginia.

BETTI NUMBERS OF 3-FOLD SYMMETRIC PRODUCTS; A CORRECTION

By M. Richardson

In a previous paper[1] the writer gave a general method for computing the Betti numbers of a complex k obtained by identifying the points of a given complex which are congruent under a finite group G of transformations subject to certain combinatorial conditions. The Betti numbers of k appeared as the ranks[2] of certain matrices (x_{ij}^m). As one of the applications of this general method, the Betti numbers of the 3-fold symmetric product k_{3n} of a given complex K_n were computed and the results published without proof.[3] In calculating the rank of the matrix (x_{ij}^m) for this case, an error was made, which invalidates the formulas of Theorem 5. The correct formulas are given here.

Theorem. *If k is the 3-fold symmetric product of K, then*

$$R_m(k) = \begin{cases} \frac{1}{6}\left[R_m(K_{3n}) + 2R_s + 3\sum_i (-1)^i R_i R_{m-2i}\right], & m = 3s, \\ \frac{1}{6}\left[R_m(K_{3n}) + 3\sum_i (-1)^i R_i R_{m-2i}\right], & m \neq 3s, \end{cases}$$

where $R_\alpha = R_\alpha(K)$, and $K_{3n} = K \times K \times K$.

Remark. If in the mechanical application of these formulas, a term happens to contain $R_\alpha(K)$, where $\alpha > n$, or $\alpha < 0$, this term is to be ignored.

We note also the following correction to the cited paper. Theorem 1 as stated is valid for $u = 1$. For $u > 1$, a slightly modified argument proves that $R_m(k, \pi^u)$ = the rank mod π of (x_{ij}) with elements reduced mod π, $(i = 1, \cdots, s;$ $j = 1, \cdots, r)$, p not divisible by π. This change does not affect later developments.

Institute for Advanced Study.

Received October 14, 1936. The existence of an error in the formulas was called to the author's attention by R. J. Walker. S. H. Kimball had previously and independently noted the error and obtained the correct formulas. The author had the advantage of seeing his results while preparing this note.

[1] *On the homology characters of symmetric products*, this Journal, vol. 1 (1935), pp. 50–69 For definitions, etc., see that paper.

[2] Loc. cit., Theorem 1, pp. 52–53.

[3] Loc. cit., Theorem 5, p. 61.

THE EXPANSION THEORY OF ORDINARY DIFFERENTIAL SYSTEMS OF THE FIRST ORDER

By Rudolph E. Langer

1. **Introduction.** It is a sufficiently curious fact that in the extensive literature of the expansion of arbitrary functions in series of characteristic solutions of ordinary linear differential systems almost no works dealing with the case of the differential systems of the first order are to be found.[1] This is the more remarkable since in this case the integration of the differential system is possible, and much of the analysis which invariably encumbers the discussions of systems of higher order is thereby obviated. To be sure, the Fourier's series, which stands as a prototype in this field, is usually associated with a differential system of the second order, and so the systems of higher order would naturally have suggested themselves as generalizations. Even for purposes of generalization, however, the system of the first order also merits attention, for because of the relative simplicity of its analysis a material generalization becomes possible in the way of a relaxation of restrictive hypotheses. This is found to be far from trivial. The theory of its expansions, as it is to be found in the literature, involves in fact a number of striking peculiarities, in which it contrasts sharply even with the expansion theories of the most closely analogous differential systems of the second order. Thus, by way of instance, one and the same expansion may be generated by an infinity of essentially distinct (i.e., non-equivalent) functions; and again the expansion of a given function, though it may converge, only rarely converges to the function immediately concerned.

It is the purpose of the present paper to present here a new expansion theory for the differential systems of the first order, one which differs from that in the literature and is believed to have advantages over it. It will be found to permit, on the one hand, of a material further relaxation of the restrictions upon the system, and to lead, on the other hand, to results which, on the whole, are much more nearly in consonance with those which obtain in the existing theories for systems of the second or higher orders. In particular, it will be found that the formal association of a function with an expansion is in a suitable sense unique, and that under quite customary conditions an expansion converges to the function with which it is formally associated. Even so, to be sure, some distinctive peculiarities persist. These will generally be recognizable, however, as inherent in the nature of the case.

Received April 30, 1937.

[1] A notable exception is M. H. Stone, *An unusual type of expansion problem*, Transactions of the American Mathematical Society, vol. 26 (1924), p. 335.

2. **The differential system.** The differential system to be considered, the general real differential system of the first order, is of the form

$$(1) \qquad y'(s) - [\rho q(s) + r(s)]y(s) = 0,$$
$$\alpha y(a) = \beta y(b).$$

The parameter, denoted by ρ, is to be free to range over all complex values. On the other hand, the variable s, the functions $q(s)$ and $r(s)$, and the constants α and β are to be real. Beyond this $q(s)$ and $r(s)$ are to be single-valued and summable in the sense of Lebesgue over the interval (a, b), and $q(s)$ is to fulfill a hypothesis of which the explicit statement is deferred to §3 below.

As applied to the differential equation of the system (1) the term "solution" will be understood to designate a function which is an indefinite Lebesgue integral and which satisfies the differential equation almost everywhere on the interval (a, b). In this sense the equation is solved by the formula

$$(2) \qquad y(s) = ce^{\int_a^s [\rho q(s) + r(s)]\, ds},$$

with c an arbitrary constant. Inasmuch as any two solutions must be linearly dependent, the differential equation determining their Wronskian to be zero, it follows that the formula (2) includes all solutions.

A solution of the differential equation which is not identically zero but which satisfies the boundary condition is to be called a *characteristic solution* of the differential system. The substitution of the form (2) into the boundary condition yields for the existence of such a characteristic solution the condition

$$\alpha = \beta e^{\rho \int_a^b q(s)\, ds + \int_a^b r(s)\, ds}.$$

In general this condition restricts ρ to an isolated set of "characteristic values". The exceptional cases are: first, that in which both α and β are zero, the condition then being obviously vacuous; second, that in which one but not both of the constants α, β is zero, the condition then being clearly impossible; and, third, that in which $\int_a^b q(s)\,ds$ is zero, the condition then being independent of ρ. To exclude these exceptional cases it will be assumed that

$$(3) \qquad \alpha\beta \int_a^b q(s)\, ds \neq 0.$$

When this hypothesis is fulfilled the differential system may be normalized as follows.

The relation

$$|\alpha| = |\beta| e^{\nu \int_a^b q(s)\, ds + \int_a^b r(s)\, ds}$$

determines the real constant ν. With it the change of variable

$$y(s) \equiv w(s) e^{\int_a^s [\nu q(s) + r(s)]\, ds},$$

transforms the differential system (1) into the form

$$w'(s) - [\rho - \nu]q(s)w(s) = 0,$$

$$w(a) = \frac{\alpha\beta}{|\alpha\beta|}\, w(b).$$

The substitutions

$$x = \frac{s - a}{b - a},$$

$$\lambda = (\rho - \nu) \int_a^b q(s)\, ds,$$

$$p(x) \equiv \frac{q(s)}{\int_a^b q(s)\, ds},$$

$$u(x) \equiv w(s)$$

complete the normalization, reducing the system (1) to the form

$$u'(x) - \lambda p(x)u(x) = 0, \qquad u(0) = u(1), \tag{4.1}$$

if $\alpha\beta > 0$, and to the form

$$u'(x) - \lambda p(x)u(x) = 0, \qquad u(0) = -u(1), \tag{4.2}$$

if $\alpha\beta < 0$. The function $p(x)$ is summable over the interval $(0, 1)$ and

$$\int_0^1 p(x)\, dx = 1. \tag{5}$$

3. **The hypothesis.** If there exists on $(0, 1)$ some sub-interval, say δ, upon which the function $p(x)$ is essentially of one sign,[2] then on this sub-interval the function

$$P(x) \equiv \int_0^x p(x)\, dx, \tag{6}$$

is monotone and, of course, continuous. The transformation

$$t = P(x) \tag{7}$$

accordingly maps the interval δ in a unique and continuous manner upon a corresponding interval, say ω, of the t-axis. If δ is taken to be directed positively,

[2] The term "essentially" will be used throughout the discussion in the sense of "almost everywhere".

ω will also be directed, positively or negatively according as on the interval δ the function $p(x)$ is essentially positive or essentially negative. Finally, on ω the transformation (7) has a unique inverse

$$x = P^{-1}(t), \tag{8}$$

in which $P^{-1}(t)$ is the Lebesgue integral of a summable function which essentially maintains its sign.[3]

The hypothesis to which the given differential system when normalized, i.e., in the form (4.1) or (4.2), is to be subjected is the following:

There shall exist on the fundamental interval (0, 1) *at least one open point-set, to be designated by* Δ, *upon which the coefficient function* $p(x)$ *fulfills the requirements*:

(i) *that it be essentially of one sign upon each of the intervals comprising the set* Δ;

(ii) *that for almost every value of* t *on the range* $0 < t < 1$ *the congruence*

$$P(x) \equiv t \pmod 1,$$

be fulfilled by some x *on* Δ, *but that for each value of* t *it be fulfilled by at most one* x *on* Δ.

The point-set Δ, being open, consists of enumerably many non-overlapping open intervals. These are to be designated by δ_j. The hypothesis (i) insures that under the transformation (7) there corresponds to each interval δ_j an interval ω_j on the t-axis. The hypothesis (ii) thereupon insures that these intervals ω_j are non-overlapping, and that the set of them as a whole would be obtainable by making a suitable sub-division of the interval $0 < t < 1$, and then at most relocating some or all of the individual sub-intervals by translating them through appropriate integral multiples of the unit distance. To avoid meaningless details it will be supposed that the designation of intervals δ_j is so made that the function $p(x)$ is essentially of opposite signs upon any two such intervals which abut, or, stated differently, that any number of abutting component intervals of the set Δ upon which the function $p(x)$ is essentially of the same sign will be regarded as comprising together one and the same interval δ_j. The intervals ω_j will be collectively designated as the point-set Ω. Since the transformation (8) exists on each individual interval ω_j, it evidently exists over the entire set Ω, and the point-sets Ω and Δ are the transforms of each other under the relation (7) or its inverse (8).

The simplest type of differential system fulfilling the hypothesis stated is that in which $p(x)$ is essentially positive throughout.[4] The point-set Δ consists in this case simply of the interval $0 < x < 1$. Because of its simplicity this type of system fails to display the peculiarities which are distinctive of the more general cases. A case which is still simple for the present discussion, though not amenable to earlier ones, is that in which the point-set Δ does not completely cover the interval (0, 1), the function $p(x)$ being essentially of one sign on Δ, but not essentially of either sign on any interval which is contained

[3] Cf. M. H. Stone, loc. cit., p. 343.

[4] This was completely treated by M. H. Stone, loc. cit.

in the complement of Δ. In this case as in the preceding one the point-set Δ is unique. In the general case, however, that is not so, the hypothesis on Δ being in fact generally satisfied by infinitely many different point-sets. This is always so when the total variation of the function $P(x)$ over those sub-intervals of $(0, 1)$ on which it is monotone exceeds the unit value.

The distinguishing features of the present hypotheses as compared with those of the earlier theory applying to the same differential systems may be summarized briefly as follows:

(i) the number of intervals constituting the set Δ is not restricted to be finite but may be enumerably infinite;

(ii) a set of intervals Δ need not completely cover the fundamental interval;

(iii) the function $p(x)$ is subject to no hypothesis other than that of summability over the point-set complementary to Δ;

(iv) the function $p(x)$ is not restricted to be bounded.

4. **The relation of "orthogonality".** Since the differential systems (4.1) and (4.2) are both explicitly integrable, their characteristic values and solutions are directly obtainable. These are found to be for the system (4.1)

$$\begin{aligned} &\lambda_0 = 0, && u_0(x) \equiv 1, \\ &\lambda_n = 2n\pi i, && u_{\pm n}(x) \equiv e^{\pm\lambda_n P(x)}, \qquad n = 1, 2, 3, \cdots, \end{aligned} \tag{9.1}$$

and for the system (4.2)

$$\begin{gathered} \lambda_n = (2n-1)\pi i, \\ u_{\pm n}(x) \equiv e^{\pm\lambda_n P(x)}, \qquad n = 1, 2, 3, \cdots. \end{gathered} \tag{9.2}$$

In either case the set of solutions as a whole has the property of satisfying the relations

$$\int_0^1 p(x) u_h(x) u_k(x)\, dx = \begin{cases} 0, & \text{if } h \neq -k, \\ 1, & \text{if } h = -k. \end{cases}$$

This property of weighted normality and orthogonality with respect to the interval determined by the points at which the boundary condition applies is, of course, wholly analogous to those which obtain for the characteristic solutions of differential systems generally, and which have been made fundamental to many deductions of the associated expansion theories. Precisely this property, however, will find no application in the present discussion, and it is principally in this point that the present theory for systems of the first order departs from that which is to be found in the literature.

The point-set Ω will easily be recognized to be of such a form that a function defined upon it may still consistently be specified to be periodic or quasi-periodic with respect to the unit distance. Such a specification moreover then extends the definition of the function to almost all values of t. It will be convenient to

reserve the notations $F_1(t)$ and $F_{-1}(t)$ to designate respectively functions whose definitions on Ω are extended by the relations

$$(10)\qquad \begin{aligned} &\text{(a)} & F_1(t+1) &\equiv F_1(t) \\ &\text{(b)} & F_{-1}(t+1) &\equiv -F_{-1}(t). \end{aligned}$$

In the case of a function of the type $F_1(t)$ the following will readily be verified, namely, that if the function is summable, then

$$(11)\qquad \sum_j \int_{|\omega_j|} F_1(t)\,dt = \int_0^1 F_1(t)\,dt,$$

the symbol $|\omega_j|$ designating the interval ω_j redirected if necessary so that its sense is positive. At the same time it is known[5] that in virtue of the transformation (7)

$$\int_{\omega_j} F_1(t)\,dt = \int_{\delta_j} p(x)F_1(P(x))\,dx, \qquad j = 1, 2, 3, \cdots,$$

and that conversely, if $p(x)f(x)$ is summable over Δ, then under the transformation (8)

$$\int_{\delta_j} p(x)f(x)\,dx = \int_{\omega_j} f(P^{-1}(t))\,dt, \qquad j = 1, 2, 3, \cdots.$$

Since in the case of each j the replacement of ω_j by $|\omega_j|$ may be compensated for by the replacement of $p(x)$ by $|p(x)|$ on δ_j, it will be seen that the relations above when summed with respect to j lead in virtue of the formula (11) to the result

$$(12)\qquad \int_0^1 F_1(t)\,dt = \int_\Delta |p(x)|\,f(x)\,dx,$$

the existence of either integral implying that of the other, and the functions involved being related thus:

$$(13)\qquad \begin{gathered} f(x) \equiv F_1(P(x)) \text{ on } \Delta, \\ F_1(t) \equiv f(P^{-1}(t)) \text{ on } \Omega, \qquad \text{and} \qquad F_1(t+1) \equiv F_1(t). \end{gathered}$$

Now whether the differential system in question be (4.1) or (4.2), it is seen from the appropriate formulas (9.1) or (9.2) that for every choice of the subscripts h and k, and as a function of t, the product $\{u_h(x)u_k(x)\}_{x=P^{-1}(t)}$ admits the unit as a period, and that

$$\int_0^1 u_h \cdot u_k\,dt = \begin{cases} 0, & \text{if } h \neq -k, \\ 1, & \text{if } h = -k. \end{cases}$$

It follows, therefore, from (12) that

$$(14)\qquad \int_\Delta |p(x)|\,u_k(x)u_h(x)\,dx = \begin{cases} 0, & \text{if } h \neq -k, \\ 1, & \text{if } h = -k. \end{cases}$$

It is upon this new relation of orthogonality that the present theory will be based.

[5] Cf. M. H. Stone, loc. cit., or E. W. Hobson, *The Theory of Functions of a Real Variable*, second edition, 1921, pp. 592–595.

5. **The expansion of an "arbitrary" function.** The expansions of a given function $f(x)$ in series of characteristic solutions differ somewhat in their details for the systems (4.1) and (4.2). Let the attention be focused first, therefore, upon the system (4.1). If $f(x)$ is a function given arbitrarily except for the restriction that $p(x)f(x)$ be summable over the point-set Δ, it may be regarded as formally associated with a series of characteristic solutions in the manner

$$f(x) \sim a_0 u_0 + \sum_{n=1}^{\infty} \{a_n u_n(x) + a_{-n} u_{-n}(x)\}. \tag{15}$$

The operations of multiplying this by $|\, p(x) \,|\, u_{-k}(x)$ and integrating term by term over Δ lead formally, and because of the relations (14), to a familiar "evaluation" of coefficients, i.e.,

$$a_k = \int_\Delta |\, p(x) \,|\, f(x) u_{-k}(x)\, dx. \tag{16}$$

With these coefficients, and with the explicit forms of the solutions as given by the formulas (9.1), the relation (15) takes the form

$$\begin{aligned} f(x) &\sim \tfrac{1}{2}A_0 + \sum_{n=1}^{\infty} \{A_n \cos 2n\pi P(x) + B_n \sin 2n\pi P(x)\}, \\ A_n &= 2\int_\Delta |\, p(x) \,|\, f(x) \cos 2n\pi P(x)\, dx, \\ B_n &= 2\int_\Delta |\, p(x) \,|\, f(x) \sin 2n\pi P(x)\, dx. \end{aligned} \tag{17}$$

Let $F_1(t)$ now designate the periodic function associated with $f(x)$ by the relation (13). The transformation (8) is found then, in virtue of the formula (12) to convert the expansion (17) into the Fourier's series for $F_1(t)$. The heuristic motivation for this deduction may now, of course, be abandoned. Whenever $f(x)$ is a function such that $p(x)f(x)$ is summable over Δ the formulas (16) associate with it a set of coefficients and thereby an expansion (15). The result deduced above may be formulated then as follows:

THEOREM. *If $p(x)f(x)$ is summable over the point-set Δ, there is associated with $f(x)$ an expansion* (15), (16) *in terms of the characteristic solutions of the differential system* (4.1). *This expansion is the transform under the relation $t = P(x)$ of the Fourier's series of the function $F_1(t)$ which is related to $f(x)$ in the manner* (13).

The system (4.2) may be similarly considered. If the expansion associated with $f(x)$ is indicated thus,

$$f(x) \sim \sum_{n=1}^{\infty} \{a_n u_n(x) + a_{-n} u_{-n}(x)\}, \tag{18}$$

the coefficients are found, in the manner above, to be again evaluated by the formulas (16). With these coefficients and the explicit forms (9.2) of the characteristic solutions the expansion takes the form

$$f(x) \sim \sum_{n=1}^{\infty} \{A_n \cos (2n-1)\pi P(x) + B_n \sin (2n-1)\pi P(x)\},$$

$$A_n = 2\int_\Delta |p(x)| f(x) \cos (2n-1)\pi P(x)\, dx, \tag{19}$$

$$B_n = 2\int_\Delta |p(x)| f(x) \sin (2n-1)\pi P(x)\, dx.$$

Let $F_{-1}(t)$ designate now the quasi-periodic function to which $f(x)$ corresponds under the transformation (8), i.e.,

$$F_{-1}(t) \equiv f(P^{-1}(t)) \text{ on } \Omega, \qquad F_{-1}(t+1) \equiv -F_{-1}(t). \tag{20}$$

Then under this transformation the expansion (19) is found to become the Fourier's series for the function $F_{-1}(t)$.

THEOREM. *If $p(x)f(x)$ is summable over the point-set Δ, there is associated with $f(x)$ an expansion* (18), (16) *in terms of the characteristic solutions of the differential system* (4.2). *This expansion is the transform under the relation $t = P(x)$ of the Fourier's series of the function $F_{-1}(t)$ which is related to $f(x)$ in the manner* (20).

6. **Conclusions.** Inasmuch as an arbitrary function and its expansion in characteristic solutions transform together into a Fourier's series and its generating function, it follows that these expansions admit of a theory which is coextensive with that of the Fourier's series. Moreover, the facts of this theory relative to convergence, summability, etc., either at points or over intervals, are evidently deducible simply by appropriate considerations of the manner in which the transformation (7), (8) affects the relative properties of the Fourier's series. In the present paper these considerations will be entered into only to the extent which will suffice to emphasize the differences between the present theory and the theory as given heretofore. The theorems chosen and given below differ sharply in these two theories, and in the present one are much more nearly in consonance with their analogues in the existing theories for differential systems of the second or higher orders.

Consider any function $f(x)$ and its expansion relative to a point-set Δ. Under the relation (7) they are transformed into the associated function $F_1(t)$ (or $F_{-1}(t)$, as the case may be) and its Fourier's series. Now if x is any point of the set Δ, then since this set is open it lies upon some sub-interval which is a neighborhood of it and which is wholly included in one of the intervals δ_j. Under the relation (7) this point x and this neighborhood correspond to a point t and a neighborhood of it. As is familiar, however, the behavior of the Fourier's series of $F_1(t)$ (or $F_{-1}(t)$) at the point t is completely determined by

the values of this function in such a neighborhood. Under the return transformation (8) this leads to the

THEOREM. *If $p(x)f(x)$ is summable over the point-set Δ, the behavior of the expansion of $f(x)$ in terms of the characteristic solutions of either of the differential systems* (4.1) *or* (4.2) *is determined at any point of Δ by the values of $f(x)$ in an arbitrarily small neighborhood of that point.*

Since the function $P(x)$ is monotone on each of the intervals δ_j, it follows that every point of Δ is possessed of some neighborhood upon which the transformation (7) is monotone. If upon such a neighborhood the function $f(x)$ is also monotone, or, more generally, of bounded variation, the same will, therefore, be true of the associated function $F_{\pm 1}(t)$ upon some interval which includes the corresponding point t. Under this condition, however, the Fourier's series is known to converge at the point t to the value $\frac{1}{2}[F_{\pm 1}(t+) + F_{\pm 1}(t-)]$, a fact which leads to the

THEOREM. *If $p(x)f(x)$ is summable over the point-set Δ, then at any point x of Δ in some neighborhood of which the function $f(x)$ is of bounded variation, the expansion of $f(x)$ in terms of the characteristic solutions of either of the differential systems* (4.1) *or* (4.2) *converges to the value*

$$\tfrac{1}{2}[f(x+) + f(x-)].$$

Consider now the case of a point x_b which does not belong to the point-set Δ, but which is a boundary point of it. Under the transformation (7) it corresponds to a point t_b which is a boundary point of Ω. The behavior of the Fourier's series of either of the functions $F_1(t)$ or $F_{-1}(t)$ at the point t_b is completely determined locally, i.e., by the values of the function in any ordinary neighborhood of the point. In virtue of their definitions (13) and (20), however, these functions are themselves determined in any ordinary neighborhood of t_b by their values on Ω, and this is immediately recognized to mean by their values on "a neighborhood of t_b in Ω", a term which by definition is to signify *that set of intervals ω_j, and parts of such, which are congruent* (mod 1) *to the parts of an ordinary neighborhood of t_b*. Under the return transformation (8) a neighborhood of t_b in Ω corresponds to "a neighborhood of x_b in Δ", a term which is to be analogously defined to designate *that set of intervals δ_j, and parts of such, upon which the function $P(x)$ takes on values which are congruent* (mod 1) *to values in a neighborhood of $P(x_b)$*. This leads to the

THEOREM. *If $f(x)$ is any function such that $p(x)f(x)$ is summable over the point-set Δ, and if x_b is any boundary point of Δ, then the behavior of the expansion of $f(x)$ in terms of the characteristic solutions of either of the differential systems* (4.1) *or* (4.2) *at x_b is determined by the values of $f(x)$ in an arbitrarily small neighborhood of x_b in Δ.*

The simplest case of a boundary point is that of a point, say x_l, which is an end point of an interval δ_j, and which has another such end point, say x_l' of an interval $\delta_{j'}$, corresponding to it under the congruence $P(x_l') \equiv P(x_l)$ (mod 1). In this case the neighborhoods of x_l in Δ, and of x_l' in Δ coincide, and if they

are sufficiently small they consist simply of the pair of intervals δ_j, $\delta_{j'}$ or parts of them which abut x_l and x_l' respectively. It is seen at once that if $f(x)$ is of bounded variation in such a neighborhood of x_l in Δ, then $F_{\pm1}(t)$ is of bounded variation in the neighborhood of t_l. The Fourier's series accordingly converges to $\frac{1}{2}[F_{\pm1}(t_l+) + F_{\pm1}(t_l-)]$, i.e., in the case of $F_1(t)$, to

$$\tfrac{1}{2}\,[\lim_{t\to t_l} F_1(t) + \lim_{t\to t_l'} F_1(t)]_{t \text{ on } \Omega},$$

and in the case of $F_{-1}(t)$ to

$$\tfrac{1}{2}\,[\lim_{t\to t_l} F_{-1}(t) + (-1)^k \lim_{t\to t_l'} F_{-1}(t)]_{t \text{ on } \Omega},$$

where k is the integer which is determined by the relation $t_l' = t_l + k$.

THEOREM. *Let $f(x)$ be any function such that $p(x)f(x)$ is summable over the point-set Δ, and let x_l and x_l' be any pair of end points of intervals of Δ which are related so that $P(x_l') = P(x_l) + k$, where k is an integer. Then if $f(x)$ is of bounded variation in an arbitrarily small neighborhood of x_l in Δ, its expansion in terms of the characteristic solutions of the differential system* (4.1) *converges at x_l to the value*

$$\tfrac{1}{2}\,[\lim_{x\to x_l} f(x) + \lim_{x\to x_l'} f(x)]_{x \text{ on } \Delta},$$

and its expansion in terms of the characteristic solutions of the differential system (4.2) *converges at x_l to the value*

$$\tfrac{1}{2}\,[\lim_{x\to x_l} f(x) + (-1)^k \lim_{x\to x_l'} f(x)]_{x \text{ on } \Delta}.$$

If Δ contains intervals which abut the points $x = 0$ and $x = 1$, these points stand in the relation of x_l and x_l' to each other, with the integer k as 1 or -1 in virtue of the relation (5). Hence we have the following

COROLLARY. *If $p(x)f(x)$ is summable over Δ, if $x = 0$ and $x = 1$ are end points of intervals of Δ, and if $f(x)$ is of bounded variation in some neighborhoods of these points, then the expansion of $f(x)$ in terms of the characteristic solutions of the differential system* (4.1) *converges at $x = 0$ and at $x = 1$ to the value*

$$\tfrac{1}{2}[f(0+) + f(1-)],$$

while the expansion of $f(x)$ in terms of the characteristic solutions of the differential system (4.2) *converges at $x = 0$ to the value*

$$\tfrac{1}{2}[f(0+) - f(1-)],$$

and converges at $x = 1$ to the value

$$\tfrac{1}{2}[f(1-) - f(0+)].$$

The function $F_1(t)$ (or $F_{-1}(t)$) will evidently be null, i.e., will vanish almost everywhere, if and only if the function $f(x)$ is null on the point-set Δ. The

condition that the difference of two functions be null, which is necessary and sufficient for the identity of all their Fourier's coefficients, leads, therefore, to the

THEOREM. *If $f_1(x)$ and $f_2(x)$ are any two functions such that $p(x)f_1(x)$ and $p(x)f_2(x)$ are summable over Δ, a necessary and sufficient condition that their expansions in terms of the characteristic solutions of either of the differential systems* (4.1) *or* (4.2) *be identical is that they coincide almost everywhere on Δ.*

Since the expansion coefficients as given by the formulas (16) involve the values of $f(x)$ only over Δ, it is evident that every expansion is strictly relative to a point-set Δ. For any specifically given function $f(x)$ there will, therefore, be as many expansions in terms of characteristic solutions of either of the differential systems (4.1) or (4.2) as there are point-sets Δ which fulfill the hypotheses stated in §3, and over which $p(x)f(x)$ is summable. Any two such point-sets Δ may, of course, have points in common, and at such points the expansions of a given function relative to the two point-sets in question will be seen at once to behave identically. For since a point common to two sets Δ is an inner point of each, it possesses a neighborhood which is also contained in each of the sets. This point and neighborhood correspond to a point on the t-axis and a neighborhood of it which is contained in each of the corresponding point-sets Ω. The functional values in this neighborhood suffice, however, for the determination of the behavior of the Fourier's series at the point in question, and the result of this determination is, therefore, independent of which point-set Ω is held in mind.

The case of a unique point-set Δ was found in §3 to be exceptional. It is clear that the case of a unique expansion for a function $f(x)$ such that $p(x)f(x)$ is summable over the interval $(0, 1)$ is exceptional in precisely the same sense. The case of such a unique expansion relative to the interval $(0, 1)$ occurs only in connection with the simplest systems, namely, those in which $p(x)$ is almost everywhere positive. In general, there is a multiplicity of expansions for each given function, and this feature must be regarded as one in which the expansion theory for differential systems of the first order as here given is at variance with the existing analogous theories for differential systems of higher orders.

UNIVERSITY OF WISCONSIN.

TRANSFORMATION OF DIFFERENTIAL EQUATIONS IN THE NEIGHBORHOOD OF SINGULAR POINTS

By C. I. Lubin

1. **Introduction.** The singular points of the system of differential equations

$$\frac{dx}{dt} = X(x, y), \qquad \frac{dy}{dt} = Y(x, y), \tag{1}$$

where the functions $X(x, y)$ and $Y(x, y)$ both vanish at a point, have been considered by many mathematicians. Beginning with Briot and Bouquet the list includes Poincaré, Picard, Bendixson, and continues with Dulac, Malmquist, Perron, Birkhoff, and many others. A full bibliography, particularly for complex differential equations, can be found in Mémorial des Sciences Mathématiques, Fascicule LXI, *Points singuliers des équations différentielles*, by H. M. Dulac.

The discussion here assumes the functions X and Y to be real and analytic in x and y, in the neighborhood of the singular point. The variables x, y, t take on only real values, and only real transformations are introduced. Furthermore, it is assumed that the singular point considered is at the origin and that the first degree terms in $X(x, y)$ and $Y(x, y)$ are such that the differential equation can be transformed by a linear transformation to the form[1]

$$\begin{aligned} \frac{dx}{dt} &= -y + \sum_{i,j=0}^{\infty} a_{ij}x^i y^j = X(x, y), \\ \frac{dy}{dt} &= x + \sum_{i,j=0}^{\infty} b_{ij}x^i y^j = Y(x, y) \qquad (i + j \geqq 2). \end{aligned} \tag{1.1}$$

At such a point the solutions are closed curves or spirals and the point is called a center or a focal point, respectively.

It is proposed here to find a canonical form for the system (1.1) and to discuss the properties of the transformation attaining that form.

2. **Failure of the usual method.** It appears from the discussion below that it is not always possible to set up formal power series

$$\begin{aligned} u &= x + \sum c_{ij}x^i y^j = f(x, y), \\ v &= y + \sum d_{ij}x^i y^j = g(x, y) \end{aligned} \tag{2}$$

Received July 15, 1936.

[1] Poincaré, Jour. de Math., (4), vol. 1, p. 172; Picard, *Traité d'Analyse*, 1928, Chap. IX.

which transform formally the system (1.1) into

$$\frac{du}{dt} = -v, \qquad \frac{dv}{dt} = u. \tag{2.1}$$

Consequently this form cannot serve as the canonical form for the system of differential equations (1.1).

If there are formal power series (2) leading to the system (2.1), the expressions $f(x, y)$ and $g(x, y)$ of (2) would formally satisfy the partial differential equations

$$\begin{aligned} X(x, y)\frac{\partial f}{\partial x} + Y(x, y)\frac{\partial f}{\partial y} &= -g, \\ X(x, y)\frac{\partial g}{\partial x} + Y(x, y)\frac{\partial g}{\partial y} &= f. \end{aligned} \tag{2.2}$$

Introduce the new variables (see Picard and Poincaré, loc. cit.)

$$x = \rho \cos \theta, \qquad y = \rho \sin \theta. \tag{2.21}$$

The partial differential equations (2.2) become

$$\begin{aligned} \frac{\partial f}{\partial \theta}\left(1 + \frac{\Omega}{\rho^2}\right) + \frac{\partial f}{\partial \rho}\frac{R}{\rho} &= -g, \\ \frac{\partial g}{\partial \theta}\left(1 + \frac{\Omega}{\rho^2}\right) + \frac{\partial g}{\partial \rho}\frac{R}{\rho} &= f, \end{aligned} \tag{2.3}$$

where Ω and R are given by

$$\begin{aligned} \Omega &= \sum_{i=3}^{\infty} \rho^i \Omega_i(\theta) = \rho(-\sin\theta \bar{X} + \cos\theta \bar{Y}), \\ R &= \sum_{i=3}^{\infty} \rho^i R_i(\theta) = \rho(\cos\theta \bar{X} + \sin\theta \bar{Y}), \end{aligned} \tag{2.31}$$

and

$$\bar{X} = X(\rho\cos\theta, \rho\sin\theta) + \rho\sin\theta = \sum_{i=2}^{\infty} \rho^i X_i(\theta),$$

$$\bar{Y} = Y(\rho\cos\theta, \rho\sin\theta) - \rho\cos\theta = \sum_{i=2}^{\infty} \rho^i Y_i(\theta).$$

In the above expressions Ω_i and R_i, and X_i and Y_i stand for homogeneous polynomials of degree i in $\sin\theta$ and $\cos\theta$, and consequently can be written as linear expressions in $\cos j\theta$, $\sin j\theta$, etc., i.e., in the form

$$\begin{aligned} &a_i \cos i\theta + a_{i-2}\cos(i-2)\theta + \cdots + b_i \sin i\theta \\ &\qquad + b_{i-2}\sin(i-2)\theta + \cdots . \end{aligned} \tag{2.32}$$

The formal expansions (2) become in terms of ρ, sin θ, and cos θ

$$(2.33) \qquad f = \rho \cos \theta + \sum_{i=2}^{\infty} \rho^i f_i(\theta), \qquad g = \rho \sin \theta + \sum_{i=2}^{\infty} \rho^i g_i(\theta),$$

where the f_i and g_i are trigonometric expressions of the type just described, (2.32). The existence of the formal expansions (2) in terms of x and y implies the existence of the formal expansions (2.33) in terms of ρ, cos θ, and sin θ, and conversely, the existence of the formal expansions (2.33) implies the existence of those of (2).

To obtain the f_i and g_i, substitute the formal series (2.33) in (2.3) and collect terms in ρ^n. We get

$$(2.4) \qquad \frac{df_n}{d\theta} = -g_n + l_n, \qquad \frac{dg_n}{d\theta} = f_n + m_n,$$

where l_n and m_n are trigonometric sums of the above type (2.32). Furthermore, the l_n and m_n involve f_i and g_i only for i less than n.

Now suppose $f_1, g_1, \cdots, f_{n-1}, g_{n-1}$ have all been computed and are of the type (2.32). To compute f_n and g_n, set up from (2.4) the differential equation

$$(2.41) \qquad \frac{d^2 f_n}{d\theta^2} = -f_n + \sum A_j^n \cos j\theta + B_j^n \sin j\theta \qquad (j = n, n-2, \cdots),$$

where the A_j^n, B_j^n arise from m_n and l_n and thus involve f_i and g_i only for i less than n.

The solution of the differential equation (2.41) is

$$(2.42) \qquad f_n = \alpha_n \cos \theta + \beta_n \sin \theta + \sum \frac{A_j^n \cos j\theta + B_j^n \sin j\theta}{1 - j^2},$$

where α_n and β_n are arbitrary and where $j = n, n-2, \cdots, 0$, or $j = n, n-2, \cdots, 1$, depending on whether n is even or odd.

When $j = 1$, a case which can only occur when n is odd, and when either $A_1^n \neq 0$ or $B_1^n \neq 0$, the solution (2.42) fails. Consequently, in general, the system of partial differential equations (2.3) does not have a formal solution in terms of ρ, sin θ, cos θ, of required type. By an expression of required type is meant a sum of the form (2.33), i.e., $\Sigma \rho^i f_i(\theta)$, where the $f_i(\theta)$ are expressible in the form (2.32). Thus the partial differential equations (2.2) do not always have a formal power series solution.

Because of this we modify the partial differential equations and consequently obtain a different normal form (v. §4 et seq.).

Let us note here the form taken by f_n, g_n of the solution, when it exists. Suppose for all n, say up to and including n', the solution (2.42) has not failed. Consider the solution for some odd number n less than n'. The expression

on the right in (2.41) then contains no term in $\sin\theta$ or $\cos\theta$, i.e., both A_1^n and B_1^n equal zero. If in (2.4)

$$l_n = a_n \cos\theta + b_n \sin\theta + \cdots,$$

we have

$$m_n = -a_n \sin\theta + b_n \cos\theta + \cdots,$$

where the terms in $\cos i\theta$ and $\sin i\theta$ for $i > 1$ have not been written.

If f_n has the form (2.42), g_n becomes

$$g_n = -\frac{df_n}{d\theta} + l_n = \alpha_n \sin\theta - \beta_n \cos\theta + a_{1n}\cos\theta + b_{1n}\sin\theta + \cdots, \tag{2.43}$$

where again only the terms in $\sin\theta$, $\cos\theta$ have been written.

3. A preliminary transformation. In order to simplify subsequent discussion a real analytic transformation will be introduced in the theorem below.

Suppose that the expressions f_i, g_i ($i = 1, 2, \cdots, 2k$) in f and g (2.33) when computed by taking all the arbitrary constants α_i, β_i zero are of required type but that the expressions for f_{2k+1}, g_{2k+1} fail to be. That is, let us suppose the differential equation (2.41), for $n = 2k + 1$, has terms in $\sin\theta$ or $\cos\theta$, i.e., A_1^n or $B_1^n \neq 0$.

By a method of Dulac,[2] we have the following

THEOREM. *By a real analytic transformation*

$$\begin{aligned} w &= x + f_2(x, y) + \cdots + f_{2k+1}(x, y) = f(x, y), \\ z &= y + g_2(x, y) + \cdots + g_{2k+1}(x, y) = g(x, y), \end{aligned} \tag{3}$$

where $f_i(x, y)$, $g_i(x, y)$ are polynomials of degree i in x and y, obtained by substituting $x = \rho\cos\theta$, $y = \rho\sin\theta$ in $\rho^i f_i(\theta)$, $\rho^i g_i(\theta)$, $(i = 2, 3, \cdots, 2k)$, of (2.42) *and* (2.43) *and in $\rho^{2k+1} f_{2k+1}(\theta)$, $\rho^{2k+1} g_{2k+1}(\theta)$ of* (3.22) *below, the differential equations* (1.1) *are reduced to*

$$\begin{aligned} \frac{dw}{dt} &= -z + Aw(w^2 + z^2)^k - Bz(w^2 + z^2)^k + \cdots = W(w, z), \\ \frac{dz}{dt} &= w + Bw(w^2 + z^2)^k + Az(w^2 + z^2)^k + \cdots = Z(w, z), \end{aligned} \tag{3.1}$$

where for the real analytic functions $W(w, z)$ and $Z(w, z)$ terms of higher degree than $(2k + 1)$ *are not written and where A and B are constants, not both zero, uniquely determined, in* (3.41) *below.*

[2] Dulac, Bull. Sci. Math., vol. 51 (1923), p. 74.

Let the differential equations causing the difficulty in the solution of the partial differential equations (2.3) be written

$$(3.2)\qquad \begin{aligned} \frac{df_{2k+1}}{d\theta} &= -g_{2k+1} + a\cos\theta + b\sin\theta + l_{2k+1},\\ \frac{dg_{2k+1}}{d\theta} &= f_{2k+1} + c\cos\theta + d\sin\theta + m_{2k+1}, \end{aligned}$$

where l_{2k+1}, m_{2k+1} contain $\sin j\theta$, $\cos j\theta$, only for $j = 3, 5, \cdots, 2k+1$, and where either $a + d \neq 0$ or $b - c \neq 0$ or both subsist.

Introduce the trigonometric functions $f^*(\theta)$ and $g^*(\theta)$, of the required type (2.32), which satisfy the differential equations

$$(3.21)\qquad \frac{df^*}{d\theta} = -g^* + l_{2k+1}, \qquad \frac{dg^*}{d\theta} = f^* + m_{2k+1},$$

where the arbitrary constants of the solution are chosen as zero. Since l_{2k+1} and m_{2k+1} contain no terms in $\sin\theta$ or $\cos\theta$, there are no such terms in f^* or g^*. Now write

$$(3.22)\qquad f_{2k+1} = f^* + \alpha\cos\theta, \qquad g_{2k+1} = g^* + \beta\cos\theta,$$

where α and β are taken as

$$\alpha = -\tfrac{1}{2}(b + c), \qquad \beta = \tfrac{1}{2}(a - d).$$

If we introduce the new variables by the substitution (3), we get as in (2.2)

$$(3.3)\qquad \frac{dw}{dt} = \frac{\partial f}{\partial x}X + \frac{\partial f}{dy}Y, \qquad \frac{dz}{dt} = \frac{\partial g}{\partial x}X + \frac{\partial g}{\partial y}Y.$$

In terms of ρ and θ, the expressions on the right side become

$$(3.31)\qquad \frac{\partial f}{\partial\theta}\left(1 + \frac{\Omega}{\rho^2}\right) + \frac{\partial f}{\partial\rho}\left(\frac{R}{\rho}\right), \qquad \frac{\partial g}{\partial\theta}\left(1 + \frac{\Omega}{\rho^2}\right) + \frac{\partial g}{\partial\rho}\left(\frac{R}{\rho}\right),$$

where Ω and R have the significance of (2.31). These quantities when expressed in terms of ρ and θ are, up to terms of degree $2k + 1$ in ρ, equal to $-f$ and g respectively. The terms in ρ^{2k+1} in each expression of (3.31) are, respectively,

$$\frac{df_{2k+1}}{d\theta} - a\cos\theta - b\sin\theta - l_{2k+1}, \qquad \frac{dg_{2k+1}}{d\theta} - c\cos\theta - d\sin\theta - m_{2k+1},$$

which, on use of (3.22), become

$$(3.4)\qquad -g_{2k+1} + A\cos\theta - B\sin\theta, \qquad f_{2k+1} + B\cos\theta + A\sin\theta,$$

where

$$(3.41)\qquad A = -\tfrac{1}{2}(a + d), \qquad B = \tfrac{1}{2}(b - c).$$

In terms of x and y the expressions (3.3) become

$$
\begin{aligned}
&\frac{\partial f}{\partial x}X + \frac{\partial f}{\partial y}Y = -g + (Ax - By)(x^2 + y^2)^k + \cdots, \\
(3.5)\qquad &\\
&\frac{\partial g}{\partial x}X + \frac{\partial g}{\partial y}Y = f + (Bx + Ay)(x^2 + y^2)^k + \cdots.
\end{aligned}
$$

The expressions in (3) can be solved for x and y in terms of z and w and substituted in (3.5), whence we get the differential equations (3.1).

4. **Second preliminary transformation.** Let us start with the system of differential equations (3.1) and modify the corresponding partial differential equations as shown in (4.3) below in order to take care of the terms in ρ^{2k+1}. The system (4.3) so introduced still presents for terms of higher degree in ρ the difficulty found in the last section for the terms in ρ^{2k+1}, so that a further transformation is useful.

It is convenient to consider this under two cases with a separate theorem in each case.

Case I. $A \neq 0$.

THEOREM. *By a real analytic transformation*

$$
\begin{aligned}
&u = w + h_2(w, z) + \cdots + h_{4k+1}(w, z) = h(w, z), \\
(4.1)\qquad &\\
&v = z + q_2(w, z) + \cdots + q_{4k+1}(w, z) = q(w, z),
\end{aligned}
$$

where $h_i(w, z)$ and $q_i(w, z)$ are polynomials of degree i obtained by the substitution $w = \rho\cos\theta$, $z = \rho\sin\theta$ in $\rho^i h_i(\theta)$ and $\rho^i q_i(\theta)$ as found below ($i = 2, 3, \cdots, 4k + 1$), *the differential equations* (3.1) *are reduced to*

$$
\begin{aligned}
&\frac{du}{dt} = -v + (Au - Bv)(u^2 + v^2)^k + Lu(u^2 + v^2)^{2k} + \cdots = U(u, v), \\
(4.2)\qquad &\\
&\frac{dv}{dt} = u + (Bu + Av)(u^2 + v^2)^k + Lv(u^2 + v^2)^{2k} + \cdots = V(u, v),
\end{aligned}
$$

where in the real analytic functions $U(u, v)$ and $V(u, v)$ terms of degree higher than $(4k + 1)$ *have not been written and where L is determined below,* (4.563).

Following the method of the last section, the partial differential equations corresponding to (2.2) are

$$
\begin{aligned}
&\frac{\partial h}{\partial w}W(w, z) + \frac{\partial h}{\partial z}Z(w, z) = -q + (Ah - Bq)(h^2 + q^2)^k, \\
(4.3)\qquad &\\
&\frac{\partial q}{\partial w}W(w, z) + \frac{\partial q}{\partial z}Z(w, z) = h + (Bh + Aq)(h^2 + q^2)^k,
\end{aligned}
$$

where $W(w, z)$ and $Z(w, z)$ have the significance of (3.1).

Write, as in (2.21), $w = \rho \cos \theta$, $z = \rho \sin \theta$, whence (4.3) become

$$(4.31)\quad \begin{aligned} \frac{\partial h}{\partial \theta}\left(1 + \frac{\Omega}{\rho^2}\right) + \frac{\partial h}{\partial \rho}\frac{R}{\rho} &= -q + (Ah - Bq)(h^2 + q^2)^k, \\ \frac{\partial q}{\partial \theta}\left(1 + \frac{\Omega}{\rho^2}\right) + \frac{\partial q}{\partial \rho}\frac{R}{\rho} &= h + (Bh + Aq)(h^2 + q^2)^k, \end{aligned}$$

where

$$\Omega = \rho[-\sin\theta\, W(\rho\cos\theta, \rho\sin\theta) + \cos\theta\, Z(\rho\cos\theta, \rho\sin\theta)] - \rho^2$$
$$= B\rho^{2k+2} + \cdots,$$
$$R = \rho[\cos\theta\, W(\rho\cos\theta, \rho\sin\theta) + \sin\theta\, Z(\rho\cos\theta, \rho\sin\theta)] = A\rho^{2k+2} + \cdots.$$

Now let us seek a solution of the partial differential equations (4.31) of the form

$$(4.311)\quad h = \rho\cos\theta + \sum \rho^i h_i(\theta), \qquad q = \rho \sin\theta + \sum \rho^i q_i(\theta).$$

Using this form for the solution we arrive at the system of ordinary differential equations similar to (2.4):

$$(4.32)\quad \frac{dh_i}{d\theta} = -q_i, \qquad \frac{dq_i}{d\theta} = h_i$$

for $1 \leqq i \leqq 2k + 1$, while for $2k + 1 < i$

$$(4.33)\quad \begin{aligned} \frac{dh_i}{d\theta} &= -q_i - B\frac{dh_i}{d\theta} - A(i - 2k)h_{i-2k} + Ah_{i-2k} - Bq_{i-2k} \\ &\quad + 2k(A\cos\theta - B\sin\theta)(\cos\theta\, h_{i-2k} + \sin\theta\, q_{i-2k}) + m_i, \\ \frac{dq_i}{d\theta} &= h_i - B\frac{dq_i}{d\theta} - A(i - 2k)q_{i-2k} + Bh_{i-2k} + Aq_{i-2k} \\ &\quad + 2k(B\cos\theta + A\sin\theta)(\cos\theta\, h_{i-2k} + \sin\theta\, q_{i-2k}) + n_i, \end{aligned}$$

where m_i and n_i involve h_j and q_j for $j < (i - 2k)$.

We shall first consider the differential equations for $i \leqq 4k$ and then consider the set for $i = 4k + 1$ separately. Let us suppose that for some integer $i = s$ ($s \leqq 4k$), the h_i and q_i for $i \leqq s - 1$ are of the required type. Furthermore let us suppose that the h_i, q_i for $i = 2, 3, \cdots, (s - 2k - 1)$ have been determined completely but that for $i = (s - 2k), \cdots, (s - 1)$ there are terms in $\sin\theta$ and $\cos\theta$ with arbitrary coefficients which are still undetermined. The solutions of these differential equations for even subscripts exist and are of the required type. For odd subscripts the solution of the differential equations for $i < s$ by virtue of (2.42) and (2.43) is

$$(4.34)\quad \begin{aligned} h_i &= M_i + \alpha_i \cos\theta + \beta_i \sin\theta, \\ q_i &= N_i + (a_i - \beta_i)\cos\theta + (b_i + \alpha_i)\sin\theta, \end{aligned}$$

where for $i \leqq s - 2k$, α_i and β_i are determined as indicated below, (4.37), but are still left undetermined for $s - 2k < i \leqq s - 1$.

From (4.32) we see that the solutions of the required type (4.34) exist for $i = s \leqq 2k + 1$ with M_i, N_i, a_i, and b_i all zero. We need then consider only the differential equations for those odd numbers s, such that $2k + 1 < s < 4k + 1$. In the expressions (4.34) take $i = s - 2k$ and substitute them in the differential equations (4.33) with $i = s$. We get

$$
(4.35)\quad \begin{aligned}
\frac{dh_s}{d\theta} &= -q_s + \cos\theta\,(\alpha_{s-2k}A(-s + 4k + 1) + a_s) \\
&\qquad + \sin\theta\,(-2kB\alpha_{s-2k} + \beta_{s-2k}A(-s + 2k + 1) + b_s) + m_s^*, \\
\frac{dq_s}{d\theta} &= h_s + \cos\theta\,(\beta_{s-2k}A(s - 2k - 1) + 2kB\alpha_{s-2k} + c_s) \\
&\qquad + \sin\theta(A\alpha_{s-2k}(-s + 4k + 1) + d_s) + n_s^*,
\end{aligned}
$$

where m_s^* and n_s^* have no terms in $\sin\theta$ or $\cos\theta$ and involve h_i and q_i only for $i \leqq s - 2k$, and where also the coefficients a_s, b_s, c_s, d_s, do the same. Note that these quantities do not depend on α_{s-2k}, β_{s-2k}.

The differential equations (4.35) lead to

$$
(4.36)\quad \begin{aligned}
\frac{d^2h_s}{d\theta^2} &= -h_s + \sin\theta\,(2\alpha_{s-2k}A(s - 4k - 1) - F_s) \\
&\qquad + \cos\theta\,(-4kB\alpha_{s-2k} - 2\beta_{s-2k}A(s - 2k - 1) + E_s) + M_s,
\end{aligned}
$$

where $E_s = b_s - c_s$, $F_s = a_s + d_s$ and $M_s = -n_s^* + dm_s^*/d\theta$.

Now choose

$$
(4.37)\quad \begin{aligned}
\alpha_{s-2k} &= \frac{F_s}{2A(s - 4k - 1)}, \\
\beta_{s-2k} &= \frac{1}{2A(s - 2k - 1)}\left[E_s - \frac{2kBF_s}{A(s - 4k - 1)}\right].
\end{aligned}
$$

This can be done for $s \neq 4k + 1$.

With this choice of constants, the differential equation (4.36) has a solution of the required type where we now choose the arbitrary coefficients α_s, β_s to be zero (i.e., when $s > 2k$). Hence there exist the required functions h_i, q_i, $i = 2k + 3, \cdots, 4k - 1$, by virtue of the proper choice of the coefficients α_{i-2k}, β_{i-2k}. These were the terms left undetermined in h_i, q_i, $i < 2k + 1$, so now with this set we have the functions $h_i, q_i, i = 1, 2, \cdots, 2k, 2k + 2, \cdots, 4k$. The arbitrary constants α_{2k+1}, β_{2k+1} are still to be determined.

To determine h_{4k+1}, q_{4k+1}, let us consider the differential equations (4.35) for $s = 4k + 1$; they become

$$
(4.4)\quad \begin{aligned}
\frac{dh_{4k+1}}{d\theta} &= -q_{4k+1} + a_{4k+1}\cos\theta + (-2kB\alpha_{2k+1} - 2Ak\beta_{2k+1} \\
&\qquad + b_{4k+1})\sin\theta + m_{4k+1}^*, \\
\frac{dq_{4k+1}}{d\theta} &= h_{4k+1} + (2kA\beta_{2k+1} + 2kB\alpha_{2k+1} + c_{4k+1})\cos\theta + d_{4k+1}\sin\theta + n_{4k+1}^*,
\end{aligned}
$$

where the terms have the significance given above (4.35).

The second order differential equation becomes

$$\frac{d^2h_{4k+1}}{d\theta^2} = -h_{4k+1} + E_{4k+1}\cos\theta - F_{4k+1}\sin\theta \tag{4.41}$$
$$-4k(B\alpha_{2k+1} + A\beta_{2k+1})\cos\theta + M_{4k+1}.$$

We see that no choice of α_{2k+1}, β_{2k+1} suffices to give (4.41) a solution of the required type, unless $(a_{4k+1} + d_{4k+1}) = 0$, i.e., $F_{4k+1} = 0$.

If $(a_{4k+1} + d_{4k+1}) = 0$, choose

$$B\alpha_{2k+1} + A\beta_{2k+1} = (b_{4k+1} - c_{4k+1})/4k = E_{4k+1}/4k, \tag{4.42}$$

and the solution h_{4k+1} of (4.41) will be of the required type, while the constant L of the lemma is zero.

However, $F_{4k+1} = 0$ is not the general case, so we proceed under the assumption $F_{4k+1} \neq 0$.

Now as in (3) use the functions h_i, q_i, $i = 2, 3, \cdots, 4k$ just found, to make the change of variable (4.1) which in terms of ρ and θ becomes

$$u = \bar{h}(\rho, \theta) = \rho\cos\theta + \rho^2 h_2(\theta) + \cdots + \rho^{4k+1} h_{4k+1}(\theta),$$
$$v = \bar{q}(\rho, \theta) = \rho\sin\theta + \rho^2 q_2(\theta) + \cdots + \rho^{4k+1} q_{4k+1}(\theta),$$

where $h_{4k+1}(\theta)$, $q_{4k+1}(\theta)$, and the coefficients α_{2k+1}, β_{2k+1}, in $h_{2k+1}(\theta)$, $q_{2k+1}(\theta)$ are determined below.

In terms of ρ and θ we have, as before:

$$\text{(4.51)}\quad \begin{aligned} &\text{(a)} \quad \frac{du}{dt} = \frac{\partial \bar{h}(\rho, \theta)}{\partial\theta}\left(1 + \frac{\Omega}{\rho^2}\right) + \frac{\partial \bar{h}(\rho, \theta)}{\partial\rho}\,\frac{R}{\rho}, \\ &\text{(b)} \quad \frac{dv}{dt} = \frac{\partial \bar{q}(\rho, \theta)}{\partial\theta}\left(1 + \frac{\Omega}{\rho^2}\right) + \frac{\partial \bar{q}(\rho, \theta)}{\partial\rho}\,\frac{R}{\rho}, \end{aligned}$$

where Ω and R have the significance of (4.31).

But up to terms of degree $4k + 1$ in ρ we have the two expressions in (4.51) equal, respectively, to

$$\text{(4.52)}\quad \begin{aligned} &\text{(a)} \quad -\bar{q} + (A\bar{h} - B\bar{q})(\bar{h}^2 + \bar{q}^2)^k, \\ &\text{(b)} \quad \bar{h} + (B\bar{h} + A\bar{q})(\bar{h}^2 + \bar{q}^2)^k. \end{aligned}$$

The terms in ρ^{4k+1} in (4.51a) and (4.51b) are

$$\text{(4.53)}\quad \begin{aligned} &\text{(a)} \quad \frac{dh_{4k+1}}{d\theta} + B(-\alpha_{2k+1}\sin\theta + \beta_{2k+1}\cos\theta) + a'\cos\theta + b'\sin\theta \\ &\qquad + A(\alpha_{2k+1}\cos\theta + \beta_{2k+1}\sin\theta)(2k + 1) + m'(\theta), \\ &\text{(b)} \quad \frac{dq_{4k+1}}{d\theta} + B(\alpha_{2k+1}\cos\theta + \beta_{2k+1}\sin\theta) + c'\cos\theta + d'\sin\theta \\ &\qquad + A(\alpha_{2k+1}\sin\theta - \beta_{2k+1}\cos\theta)(2k + 1) + n'(\theta), \end{aligned}$$

while the terms in ρ^{4k+1} in (4.52a) and (4.52b) are

$$
(4.54)\quad
\begin{aligned}
&\text{(a)}\ -q_{4k+1} + A(\alpha_{2k+1}\cos\theta + \beta_{2k+1}\sin\theta) + a''\cos\theta + b''\sin\theta \\
&\qquad -B(\alpha_{2k+1}\sin\theta - \beta_{2k+1}\cos\theta) + 2k(A\cos\theta - B\sin\theta)\alpha_{2k+1} + m''(\theta), \\
&\text{(b)}\ h_{4k+1} + \beta_{2k+1}(-A\cos\theta + B\sin\theta) + c''\cos\theta + d''\sin\theta \\
&\qquad\qquad + \alpha_{2k+1}(B\cos\theta + A\sin\theta)(2k+1) + n''(\theta),
\end{aligned}
$$

where $\sin\theta$, $\cos\theta$, α_{2k+1}, β_{2k+1} occur only where explicitly written.

Now introduce the functions $h^*(\theta)$ and $q^*(\theta)$ satisfying the differential equations

$$
(4.55)\qquad \frac{dh^*}{d\theta} = -q^* + \overset{*}{m}_{4k+1}(\theta), \qquad \frac{dq^*}{d\theta} = h^* + \overset{*}{n}_{4k+1}(\theta),
$$

where

$$
\begin{aligned}
m''(\theta) - m'(\theta) &= \overset{*}{m}_{4k+1}(\theta), \\
n''(\theta) - n'(\theta) &= \overset{*}{n}_{4k+1}(\theta),
\end{aligned}
$$

as written in (4.4). This system (4.55) has a solution $h^*(\theta)$, $q^*(\theta)$ of the required type.

Now, for $h_{4k+1}(\theta)$, $q_{4k+1}(\theta)$ of (4.5) write

$$
\begin{aligned}
h_{4k+1}(\theta) &= h^*(\theta) + \mu\sin\theta, \\
q_{4k+1}(\theta) &= q^*(\theta) + \lambda\sin\theta,
\end{aligned}
$$

and substitute in (4.53a). This becomes

$$
(4.56)\quad
\begin{aligned}
&\frac{dh^*}{d\theta} + \mu\cos\theta + \alpha_{2k+1}[-B\sin\theta + (2k+1)A\cos\theta] + a'\cos\theta \\
&\quad + b'\sin\theta + \beta_{2k+1}[B\cos\theta + (2k+1)A\sin\theta] + m'(\theta) \\
&\quad = -q^* - \lambda\sin\theta + \lambda\sin\theta + \overset{*}{m}_{4k+1} + \mu\cos\theta \\
&\quad + \beta_{2k+1}[B\cos\theta + (2k+1)A\sin\theta] + a'\cos\theta + b'\sin\theta \\
&\quad + \alpha_{2k+1}[-B\sin\theta + (2k+1)A\cos\theta] + m'(\theta) = -q_{4k+1} + m''(\theta) \\
&\quad + a''\cos\theta + b''\sin\theta + A(\alpha_{2k+1}\cos\theta + \beta_{2k+1}\sin\theta) \\
&\quad - B(\alpha_{2k+1}\sin\theta - \beta_{2k+1}\cos\theta) + 2k\alpha_{2k+1}(A\cos\theta - B\sin\theta) \\
&\quad + \cos\theta(-a_{4k+1} + \mu) + (-b_{4k+1} + 2kB\alpha_{2k+1} + 2kA\beta_{2k+1} + \lambda)\sin\theta;
\end{aligned}
$$

i.e., (4.53a) becomes equal to (4.54a) except for the expression

$$
(4.561)\qquad (-a_{4k+1} + \mu)\cos\theta + (-b_{4k+1} + 2kB\alpha_{2k+1} + 2kA\beta_{2k+1} + \lambda)\sin\theta.
$$

Hence the expression (4.51a) equals the expression (4.52a) up to and including terms of degree $4k + 1$ in ρ except for (4.561).

In the same way it is found that the expression (4.51b) equals the expression (4.52b) up to and including terms of degree $4k + 1$ in ρ except for the terms

$$
(4.562)\qquad (\lambda - c_{4k+1} - 2kA\beta_{2k+1} - 2kB\alpha_{2k+1})\cos\theta + (-\mu - d_{4k+1})\sin\theta.
$$

Now choose the constants as follows:

$$B\alpha_{2k+1} + A\beta_{2k+1} = (b_{4k+1} - c_{4k+1})/4k,$$
$$\mu = \tfrac{1}{2}(-a_{4k+1} - d_{4k+1}),$$
$$\lambda = \tfrac{1}{2}(b_{4k+1} + c_{4k+1}),$$

whence (4.561) and (4.562) become $L \cos \theta$ and $L \sin \theta$ respectively, where

$$L = \tfrac{1}{2}(-d_{4k+1} - a_{4k+1}). \tag{4.563}$$

The theorem follows now as in the preceding section.

Case II. $A = 0, B \neq 0$.

The partial differential equations (4.3) now become

$$\begin{aligned} \frac{\partial h}{\partial \theta}\left(1 + \frac{\Omega}{\rho^2}\right) + \frac{\partial h}{\partial \rho}\frac{R}{\rho} &= -q - Bq(h^2 + q^2)^k, \\ \frac{\partial q}{\partial \theta}\left(1 + \frac{\Omega}{\rho^2}\right) + \frac{\partial q}{\partial \rho}\frac{R}{\rho} &= h + Bh(h^2 + q^2)^k. \end{aligned} \tag{4.6}$$

Again let the formal solution be (4.311) and let us follow the procedure of Case I. If we attempt now to solve the differential equations corresponding to (4.35) in h_s, q_s, we are led to

$$\frac{d^2 h_s}{d\theta^2} = -h_s + F_s \sin \theta + (-4kB\alpha_{s-2k} + E_s)\cos \theta + M_s \tag{4.61}$$

where the terms here have the significance of (4.36).

We thus see that a solution of the required type exists provided $F_s = 0$ and provided the arbitrary constant is chosen as follows:

$$\alpha_{s-2k} = E_s/4kB.$$

However in the general case, for some $s = 2p + 1$,

$$F_{2p+1} \neq 0, \qquad p > k,$$

in which case we have the following theorem:

Theorem. *By a real analytic transformation*

$$u = w + \cdots,$$
$$v = z + \cdots,$$

the differential equations (3.1) *are reduced to*

$$\begin{aligned} \frac{du}{dt} &= -v - Bv(u^2 + v^2)^k + Au(u^2 + v^2)^p + \cdots, \\ \frac{dv}{dt} &= u + Bu(u^2 + v^2)^k + Av(u^2 + v^2)^p + \cdots. \end{aligned} \tag{4.7}$$

The discussion here is much the same as in the preceding lemmas and will not be repeated.

Let us note a further change which can be introduced both in (4.2) and (4.7). It is essentially a change in the parameter t such that

$$dt = d\tau/\{1 + B(u^2 + v^2)^k\}. \tag{4.71}$$

The system (4.2) becomes

$$\begin{aligned} \frac{du}{d\tau} &= -v + Au(u^2 + v^2)^k + L'u(u^2 + v^2)^{2k} + \cdots, \\ \frac{dv}{d\tau} &= u + Av(u^2 + v^2)^k + L'v(u^2 + v^2)^{2k} + \cdots, \end{aligned} \tag{4.72}$$

where $L' = L - AB$. The system (4.72) is the form sought under the preliminary transformations.

However, the system (4.7) under (4.71) becomes

$$\frac{du}{d\tau} = -v + Au(u^2 + v^2)^p + \cdots,$$

$$\frac{dv}{d\tau} = u + Av(u^2 + v^2)^p + \cdots,$$

which is in the form (3.1) with $A \neq 0$ so that Case I of the theorem in this section applies. Thus the ultimate form in each case is (4.72) above.

Note that after the proper change of parameter, above, has been found, a corresponding one could be made in the initial set of differential equations (1.1) at once. Beginning with the set (1.1) so modified no further change in parameter is necessary and we arrive at (4.72) under changes in the dependent variables only.

The set of differential equations (1.1) transform under the 1-1 analytic transformations (3) and (4.1) which contain arbitrary constants into a system of the above form with unique k, A, B, and L. The system so obtained is spoken of as unique. When in addition the change of parameter (4.71) is introduced, the system attained (4.72) again has a unique set of constants, $\bar{k}$, $\bar{A}$, $\bar{L}$.

These statements can be established by observing that if there were a different system attained, there would exist a 1-1 analytic transformation taking the one system into the other. This leads to a contradiction.

5. **Convergent case.** It is of interest to examine what occurs when $A_1^n \equiv 0$. For the case where the functions $X(x, y)$ and $Y(x, y)$ are polynomials, Poincaré[3] showed that if the partial differential equation

$$\frac{\partial f}{\partial \theta}\left(1 + \frac{\Omega}{\rho^2}\right) + \frac{\partial f}{\partial \rho}\frac{R}{\rho} = 0 \tag{5.1}$$

[3] Poincaré, loc. cit.

has a formal solution of the required type, that solution is also convergent. The existence of this integral of (5.1) implies the existence of analytic solutions of (2.2)[4] so that for these cases there exist analytic functions which accomplish the transformation.

This case can arise in (2.2) or (4.7). In the former the system of differential equations becomes

$$\frac{du}{dt} = -v, \qquad \frac{dv}{dt} = u.$$

In the case of (4.7) the system becomes

$$\frac{du}{dt} = -v(1 + B(u^2 + v^2)^k),$$

$$\frac{dv}{dt} = u(1 + B(u^2 + v^2)^k).$$

6. **The formal transformation.** If we attempt now to transform the system of differential equations (4.2) or (4.72), we find there does exist a formal transformation of the required type. The method followed in establishing this is precisely that of §4; in fact, the partial differential equations here

$$\frac{\partial h}{\partial \theta}\left(1 + \frac{\Omega}{\rho^2}\right) + \frac{\partial h}{\partial \rho}\frac{R}{\rho} = -q + (Ah - Bq)(h^2 + q^2)^k + Lh(h^2 + q^2)^{2k},$$

$$\frac{\partial q}{\partial \theta}\left(1 + \frac{\Omega}{\rho^2}\right) + \frac{\partial q}{\partial \rho}\frac{R}{\rho} = h + (Bh + Aq)(h^2 + q^2)^k + Lq(h^2 + q^2)^{2k}$$

differ but slightly from those of (4.31) and the related expressions corresponding to (4.35), (4.36), and (4.37) likewise differ very slightly from them. However the term in ρ^{4k+1} which caused the difficulty before is now taken care of by the term involving the constant L. This term also contributes the slight differences between the expressions here and the corresponding ones of §4. The result is stated in the following theorem, whose proof will not be given:

THEOREM. *The system of real differential equations*

$$(6.1)\qquad \begin{aligned} \frac{du}{dt} &= -v + (Au - Bv)(u^2 + v^2)^k + Lu(u^2 + v^2)^{2k} + \cdots = U(u, v),\\ \frac{dv}{dt} &= u + (Bu + Av)(u^2 + v^2)^k + Lv(u^2 + v^2)^{2k} + \cdots = V(u, v), \end{aligned}$$

where U and V are the real analytic functions of (4.2), *can be transformed to the unique (see* §4) *system*

$$(6.2)\qquad \begin{aligned} \frac{du_1}{dt} &= -v_1 + (Au_1 - Bv_1)(u_1^2 + v_1^2)^k + Lu_1(u_1^2 + v_1^2)^{2k},\\ \frac{dv_1}{dt} &= u_1 + (Bu_1 + Av_1)(u_1^2 + v_1^2)^k + Lv_1(u_1^2 + v_1^2)^{2k} \end{aligned}$$

[4] Birkhoff, Am. Jour. Math., vol. 49 (1927), p. 37.

by the formal transformation

$$(6.3)\qquad \begin{aligned} u_1 &= u + (\alpha_{2k+1}u + \beta_{2k+1}v)(u^2 + v^2)^k + \sum a_{ij}u^i v^j, \\ v_1 &= v + (-\beta_{2k+1}u + \alpha_{2k+1}v)(u^2 + v^2)^k + \sum b_{ij}u^i v^j, \end{aligned}$$

where

$$i + j > 2k + 1$$

$$A\beta_{2k+1} + B\alpha_{2k+1} = 0.$$

Note that, if desired, the quantity B can be taken equal to zero by virtue of the form (4.72), and here also there is but one form for the transformed differential equations.

7. **Auxiliary functions.** The series (6.3) of the last section, which accomplish the desired transformations, are formal and not necessarily convergent. Consequently these expressions are not satisfactory and we seek more appropriate ones.

Following Birkhoff,[5] introduce the associated series

$$(7.1)\qquad \begin{aligned} f(u, v) &= u + (\alpha u + \beta v)(u^2 + v^2)^k + \sum a_{ij}\left(1 - e^{\frac{-1}{[1+|a_{ij}|](u^2+v^2)^k}}\right) u^i v^j, \\ g(u, v) &= v + (-\beta u + \alpha v)(u^2 + v^2)^k + \sum b_{ij}\left(1 - e^{\frac{-1}{[1+|b_{ij}|](u^2+v^2)^k}}\right) u^i v^j. \end{aligned}$$

These series are of a type considered by Ritt[6] and by Birkhoff,[7] so that only an outline of the proof of the following theorem, which is completely analogous to a theorem of Birkhoff's (loc. cit.), is given here.

THEOREM. *The real functions $f(u, v)$, $g(u, v)$ defined in* (7.1) *are analytic in the real variables u and v for*

$$0 < \nu \leqq u^2 + v^2 \leqq \rho_0^2$$

and C^∞ for

$$0 \leqq u^2 + v^2 \leqq \rho_0^2$$

and such that

$$\left.\frac{\partial^{i+j} f}{\partial u^i \partial v^j}\right|_0 = a_{ij}\, i!\, j!, \qquad \left.\frac{\partial^{i+j} g}{\partial u^i \partial v^j}\right|_0 = b_{ij}\, i!\, j!.$$

In the region of complex space,

$$|u^2 + v^2| \leqq \rho_0^2,$$

$$|\arg u| < \frac{\pi}{8k}, \qquad |\arg v| < \frac{\pi}{8k},$$

or

$$|\pi - \arg u| < \frac{\pi}{8k}, \qquad |\pi - \arg v| < \frac{\pi}{8k},$$

[5] Birkhoff, Sitzungsberichte Preus. Akad. (1929), pp. 171–183.
[6] J. F. Ritt, Annals of Mathematics, (2), vol. 18 (1916), p. 18.
[7] Birkhoff, loc. cit.

we find

$$\left| a_{ij} u^i v^j \left(1 - e^{\frac{-1}{B_{ij}}}\right) \right| < \frac{2 | u |^i | v |^j}{| u^2 + v^2 |^k},$$

where

$$B_{ij} = \{1 + | a_{ij} |\}(u^2 + v^2)^k.$$

We also have

$$| u^2 + v^2 | \geqq \sqrt{2} \rho_1 \rho_2 ,$$

where

$$u = \rho_1 e^{i\theta_1}, \qquad v = \rho_2 e^{i\theta_2}.$$

For the terms under consideration, $i + j > 2k$, divide the treatment into the following three cases:

(a) $\qquad i \geqq k, \qquad j \geqq k;$

then

$$\frac{2 | u |^i | v |^j}{| u^2 + v^2 |^k} \leqq \frac{2 | u |^i | v |^j}{| u |^k | v |^k 2^{k/2}} < 2 | u |^{i-k} | v |^{j-k}.$$

(b) $\qquad i < k, \qquad j > 2k - i;$

then

$$\frac{2 | u |^i | v |^j}{| u^2 + v^2 |^k} \leqq \frac{2 | v |^{j-i}}{2^{i/2} | u^2 + v^2 |^{k-i}} < 2 | v |^{i+j-2k}.$$

(c) $\qquad j < k, \qquad i > 2k - j.$

As in (b), these terms are less than $2 | u |^{i+j-2k}$.

Thus we have as a dominating series the expression

$$2 \sum_{\substack{i=k \\ j=k}}^{i,j=\infty} | u |^{i-k} | v |^{j-k} + 2 \sum_{\substack{j=1 \\ i=k+1}}^{\substack{i=\infty \\ j=k}} | u |^{i+j-2k} + 2 \sum_{\substack{i=1 \\ j=k+1}}^{\substack{i=k \\ j=\infty}} | v |^{j+i-2k}.$$

Consequently we have the analyticity of the function in the open region and continuity at the origin.

Let us now consider the partial derivative $\partial/\partial u$ of the series. A typical term for this differentiation becomes

$$i a_{ij} u^{i-1} v^j \left(1 - e^{\frac{-1}{B_{ij}}}\right) - \frac{a_{ij}}{1 + | a_{ij} |} u^{i+1} v^j e^{\frac{-1}{B_{ij}}} \frac{2k}{(u^2 + v^2)^{k+1}}.$$

As above, we have

$$\left| i a_{ij} u^{i-1} v^j \left(1 - e^{\frac{-1}{B_{ij}}}\right) \right| \leqq 2i | u |^{i-k-1} | v |^{j-k}.$$

For the second part of the derivative of the term we find

$$\left| \frac{a_{ij}}{1 + |a_{ij}|} u^i v^j e^{\frac{-1}{B_{ij}}} \frac{2k}{(u^2 + v^2)^{k+1}} \right| \leqq 2k |u|^{i-k} |v|^{j-k-1},$$

provided $\qquad i > k + 1, \qquad\qquad j > k + 1.$

The terms for which these inequalities are not satisfied are treated as under (b) and (c) above. Thus again we have a dominating series and consequently the derived series converges uniformly and absolutely and hence this partial derivative has the stated properties of continuity and analyticity.

Higher derivatives are treated in the same manner.

The limits approached by the derived series are as stated. The proof merely depends on the observation that each term of the derived series $\partial^{i+j}/(\partial u^i \partial v^j)$ approaches zero with $(u^2 + v^2)$ except the one of the form

$$i!\,j!\,a_{ij}\left(1 - e^{\frac{-1}{B_{ij}}}\right),$$

whose limit is $i!\,j!\,a_{ij}$.

Introduce the variables ρ, θ in the functions (7.1), where

$$u = \rho \cos \theta, \qquad\qquad v = \rho \sin \theta.$$

The functions become

$$\begin{aligned}
\bar{f}(\rho, \theta) &= \rho \cos \theta + (\alpha \cos \theta + \beta \sin \theta)\rho^{2k+1} \\
&\qquad + \sum \rho^{i+j} a_{ij}\left(1 - e^{\frac{-1}{A_{ij}}}\right) \cos^i \theta \sin^j \theta, \\
\bar{g}(\rho, \theta) &= \rho \sin \theta + (\alpha \sin \theta - \beta \cos \theta)\rho^{2k+1} \\
&\qquad + \sum \rho^{i+j} b_{ij}\left(1 - e^{\frac{-1}{B_{ij}}}\right) \cos^i \theta \sin^j \theta,
\end{aligned} \tag{7.2}$$

where

$$A_{ij} = \{1 + |a_{ij}|\}\rho^{2k}, \qquad\qquad B_{ij} = \{1 + |b_{ij}|\}\rho^{2k},$$

which are analytic in the open region

$$0 < \rho \leqq \rho_0,$$
$$0 \leqq \theta \leqq 2\pi$$

and of class C^∞ in the closed region

$$0 \leqq \rho \leqq \rho_0,$$
$$0 \leqq \theta \leqq 2\pi.$$

Considered as power series in $\sin \theta$, $\cos \theta$, the expressions (7.2) are uniformly and absolutely convergent in the closed region. Furthermore, these functions are characterized by the following theorem:

THEOREM. *The functions* (7.2) *can be expressed as power series in* θ:

$$(7.21)\qquad \begin{aligned} \bar{f}(\rho, \theta) &= f_0(\rho) + f_1(\rho)\theta + f_2(\rho)\theta^2 + \cdots, \\ \bar{g}(\rho, \theta) &= g_0(\rho) + g_1(\rho)\theta + g_2(\rho)\theta^2 + \cdots \end{aligned}$$

convergent for

$$0 \leqq \theta \leqq 2\pi, \qquad\qquad 0 \leqq \rho \leqq \rho_0,$$

where the functions $f_i(\rho)$, $g_i(\rho)$ *are analytic for* $0 < \rho \leqq \rho_0$ *and of class* C^∞ *for* $0 \leqq \rho \leqq \rho_0$.

Consider the first expression in (7.2); we have, as seen above,

$$\left| a_{ij}\rho^{i+j}\left(1 - e^{\frac{-1}{A_{ij}}}\right) \right| < 2\rho^{i+j-2k}.$$

Consequently the series

$$\sum d_l \rho^{i+j-2k}\theta^l = 2\rho^{i+j-2k} \cosh^i\theta \sinh^j\theta$$

dominates the series obtained by expanding the term

$$a_{ij}\rho^{i+j}\left(1 - e^{\frac{-1}{A_{ij}}}\right) \cos^i\theta \sin^j\theta$$

in powers of θ.

Furthermore, the series

$$a_0(\rho) + a_1(\rho)\theta + a_2(\rho)\theta^2 + \cdots = \sum 2\rho^{i+j-2k} \cosh^i\theta \sinh^j\theta$$

converges uniformly and absolutely for $\rho \leqq \rho_1$ sufficiently small and $0 \leqq \theta \leqq 2\pi$ and all the $a_i(\rho)$ are positive. Then, since

$$|f_i(\rho)| < a_i(\rho), \qquad\qquad |g_i(\rho)| < a_i(\rho),$$

the series (7.21) converge. It is readily shown that the $f_i(\rho)$, $g_i(\rho)$ have the additional properties stated in the theorem.

A similar discussion holds for the partial derivatives, and we have the second[8]

THEOREM. *The first partial derivatives of the functions* $\bar{f}(\rho, \theta)$, $\bar{g}(\rho, \theta)$ *exist and can be expressed as the convergent series made up of the corresponding partial derivatives of the terms of the series* (7.21).

Now introduce new variables $\bar{u}$, $\bar{v}$ by means of (7.1), as follows:

$$(7.3)\qquad \bar{u} = f(u, v), \qquad\qquad \bar{v} = g(u, v).$$

The variables u, v can be expressed as

$$(7.31)\qquad u = f^*(\bar{u}, \bar{v}), \qquad\qquad v = g^*(\bar{u}, \bar{v}),$$

where the functions f^*, g^* have the same properties as $f(u, v)$ and $g(u, v)$ of the above theorem.

[8] See Bôcher, *On semi-analytic functions of two variables*, Annals of Mathematics, vol. 12 (1910–11), p. 18.

Put $\bar{u} = r \cos \phi$, $\bar{v} = r \sin \phi$ and use (7.2). Then we have

$$\begin{aligned} r \cos \phi &= \bar{f}(\rho, \theta), \\ r \sin \phi &= \bar{g}(\rho, \theta). \end{aligned} \tag{7.32}$$

The properties of the inverse transformation, the expressions for ρ, θ or ρ, $\sin \theta$ (or $\cos \theta$), can be established in the following theorems. (The proofs are omitted.)

THEOREM. *The system of equations* (7.32) *can be solved for* ρ, θ *in the form*

$$\theta = \phi + \Sigma \bar{h}_i(r)\phi^i, \qquad \rho = r + \Sigma \bar{m}_i(r)\phi^i \tag{7.4}$$

convergent for $0 \leqq \phi \leqq 2\pi$, $0 \leqq r \leqq r_0$, *where the* $\bar{h}_i(r)$ *and* $\bar{m}_i(r)$ *have the properties of* $f_i(\rho)$, $g_i(\rho)$ *in* (7.21) *above.*

For the second theorem we have:

THEOREM. *The system of equations* (7.32) *can be solved for* ρ *and* $\sin \theta$ *or* ρ *and* $\cos \theta$ *in the form*

$$\begin{aligned} \rho &= r + \sum h_{ij}(r) \cos^i \phi \sin^j \phi, \\ \sin \theta &= \sin \phi + \sum \eta_{ij}(r) \cos^i \phi \sin^j \phi, \end{aligned} \tag{7.41}$$

which are uniformly and absolutely convergent considered as power series in $\sin \phi$, $\cos \phi$, *for*

$$0 \leqq \phi \leqq 2\pi, \qquad 0 \leqq r \leqq r_0,$$

where $h_{ij}(r)$ *and* $\eta_{ij}(r)$ *are analytic in* r *for* $0 < r \leqq r_0$ *and of class* C^∞ *for* $0 \leqq r \leqq r_0$, *and are such that*

$$h_{ij}(0) = 0, \qquad \eta_{ij}(0) = 0.$$

8. **Auxiliary transformation.** In the system of differential equations (6.1) put

$$u = \rho \cos \theta, \qquad v = \rho \sin \theta,$$

and then introduce the new variables r, ϕ by means of the transformation (7.32)

$$r \cos \phi = \bar{f}(\rho, \theta), \qquad r \sin \phi = \bar{g}(\rho, \theta), \tag{8.1}$$

whence

$$r = \sqrt{\bar{f}^2 + \bar{g}^2} = M(\rho, \theta). \tag{8.11}$$

The system becomes

$$\begin{aligned} dr/dt &= Ar^{2k+1} + Lr^{4k+1} + R^*(r, \phi), \\ d\phi/dt &= 1 + Br^{2k} + \Omega^*(r, \phi), \end{aligned} \tag{8.2}$$

where the properties of the functions $R^*(r, \phi)$ and $\Omega^*(r, \phi)$ are given in the following theorems.

THEOREM. *The function $R^*(r, \phi)$ of (8.2) is analytic in r, ϕ for $0 < r \leqq r_0$, $0 \leqq \phi \leqq 2\pi$, of class C^∞ for $0 \leqq r \leqq r_0$, $0 \leqq \phi \leqq 2\pi$, and can be written $R^*(r, \phi) = R_0(r) + R_1(r)\phi + \cdots$, which converges for $0 \leqq r \leqq r_0$, $0 \leqq \phi \leqq 2\pi$.*

To prove the theorem, let us note the form of $R^*(r, \phi)$. Introducing r by means of (7.41), we get

$$\frac{dr}{dt} = \frac{\partial M}{\partial \theta}\left(1 + \frac{\Omega(\rho, \theta)}{\rho^2}\right) + \frac{\partial M}{\partial \rho}\left(\frac{R(\rho, \theta)}{\rho}\right), \tag{8.3}$$

an expression in ρ and θ, where Ω and R arise from U and V here, as they do from X and Y in (2.31). In terms of r, ϕ this becomes

$$Ar^{2k+1} + Lr^{4k+1} + R^*(r, \phi), \tag{8.31}$$

where the expression $R^*(r, \phi)$ can be written in terms of ρ and θ as

$$\left(1 + \frac{\Omega}{\rho^2}\right)\frac{\partial M}{\partial \theta} + \frac{R}{\rho}\frac{\partial M}{\partial \rho} - AM^{2k+1} - LM^{4k+1} = \bar{R}^*(\rho, \theta), \tag{8.32}$$

whence from the theorems of §7, this expression in terms of ρ and θ has the required properties. Consequently, it also has the required properties in terms of r and ϕ. Note that where ϕ enters in $R^*(r, \phi)$, it can be expressed entirely as powers of $\sin \phi$ and $\cos \phi$.

A similar result holds for $\Omega^*(r, \phi)$.

A second theorem is as follows:

THEOREM. *For any positive integer p,*

$$\lim_{r=0} \frac{R^*(r, \phi)}{r^p} = 0.$$

For the expression in (8.11), write

$$M(\rho, \theta) = M'(\rho, \theta) + M''(\rho, \theta), \tag{8.4}$$

where $M'(\rho, \theta)$ is made up only of those terms of $M(\rho, \theta)$ which involve the terms of (7.2) with coefficients a_{ij}, b_{ij} such that $i + j \leqq p$, while $M''(\rho, \theta)$ is made up of the rest of the expansion of $M(\rho, \theta)$.

Now put

$$M'(\rho, \theta) = M_1(\rho, \theta) + M_2(\rho, \theta), \tag{8.41}$$

where $M_1(\rho, \theta)$ is of the form

$$\sum \bar{a}_{ij} \cos^i\theta \sin^j\theta\, \rho^{i+j} \qquad (i + j \leqq p),$$

while $M_2(\rho, \theta)$ is made up of terms, each of which contains at least one factor of the type $e^{\frac{-1}{A l m}}$.

We have

$$\frac{\partial M_1}{\partial \theta}\left(1 + \frac{\Omega}{\rho^2}\right) + \frac{\partial M_1}{\partial \rho}\frac{R}{\rho} = AM_1^{2k+1} + LM_1^{4k+1} + N(\rho, \theta), \tag{8.42}$$

where $N(\rho, \theta)$ is a power series in ρ, $\cos\theta$, $\sin\theta$, such that each term is of degree greater than p in ρ.

Now, for the function $\bar{R}^*(\rho, \theta)$ of (8.32) we have

$$\begin{aligned}
\bar{R}^*(\rho, \theta) &= \frac{\partial}{\partial\theta}(M_1 + M_2 + M'')\left(1 + \frac{\Omega}{\rho^2}\right) + \frac{\partial}{\partial\rho}(M_1 + M_2 + M'')\frac{R}{\rho} \\
&\qquad - A(M_1 + M_2 + M'')^{2k+1} - L(M_1 + M_2 + M'')^{4k+1} \\
&= N(\rho, \theta) + AM_1^{2k+1} + LM_1^{4k+1} \qquad (8.5)\\
&\qquad - A(M_1 + M_2 + M'')^{2k+1} - L(M_1 + M_2 + M'')^{4k+1} \\
&\qquad + \frac{\partial}{\partial\theta}(M_2 + M'')\left(1 + \frac{\Omega}{\rho^2}\right) + \frac{\partial}{\partial\rho}(M_2 + M'')\frac{R}{\rho}.
\end{aligned}$$

Consequently, we see

$$\lim_{\rho=0} \frac{\bar{R}^*(\rho, \theta)}{\rho^p} = 0,$$

and thus

$$\lim_{r=0} \frac{R^*(r, \phi)}{r^p} = 0.$$

A similar result holds for $\Omega^*(r, \phi)$.

Finally introduce into the differential equations the variables $\bar{u}$, $\bar{v}$ of §7. We immediately have the following facts:

Theorem. (1) *Under the transformation* (7.3) *the system of differential equations becomes*

$$(8.6)\qquad \begin{aligned}
\frac{d\bar{u}}{dt} &= -\bar{v} + (A\bar{u} + B\bar{v})(\bar{u}^2 + \bar{v}^2)^k + L\bar{u}(\bar{u}^2 + \bar{v}^2)^{2k} + \bar{U}(\bar{u}, \bar{v}), \\
\frac{d\bar{v}}{dt} &= \bar{u} + (B\bar{u} + A\bar{v})(\bar{u}^2 + \bar{v}^2)^k + L\bar{v}(\bar{u}^2 + \bar{v}^2)^{2k} + \bar{V}(\bar{u}, \bar{v}).
\end{aligned}$$

(2) *The functions* $\bar{U}(\bar{u}, \bar{v})$, $\bar{V}(\bar{u}, \bar{v})$ *are analytic in the open region* $0 < (\bar{u}^2 + \bar{v}^2)^k \leq \rho_1^2$ *and class* C^∞ *in the closed region.*

(3) *For any integer* p,

$$\lim_{\substack{\bar{u}=0\\ \bar{v}=0}} \frac{\bar{U}(\bar{u}, \bar{v})}{(\bar{u}^2 + \bar{v}^2)^p} = 0,$$

$$\lim_{\substack{\bar{u}=0\\ \bar{v}=0}} \frac{V(\bar{u}, \bar{v})}{(\bar{u}^2 + \bar{v}^2)^p} = 0.$$

9. **Final transformation.** The final transformation is introduced in the following

Theorem. *There exist functions*

$$(9.1) \qquad f_1 = \bar{u} + \cdots, \qquad g_1 = \bar{v} + \cdots$$

analytic in the open region $\bar{u}^2 + \bar{v}^2 > 0$, *of class* C^∞ *in the closed region* $\bar{u}^2 + \bar{v}^2 \geqq 0$ *such that the differential equations* (8.6), *under the transformation* $w = f_1$, $z = g_1$, *become*

$$(9.2) \qquad \begin{aligned} \frac{dw}{dt} &= -z + (Aw - Bz)(z^2 + w^2)^k + Lw(z^2 + w^2)^{2k}, \\ \frac{dz}{dt} &= w + (Bw + Az)(z^2 + w^2)^k + Lz(z^2 + w^2)^{2k}. \end{aligned}$$

To establish this theorem let us first prove the lemma below:

Lemma. *The partial differential equation*

$$(9.3) \qquad \frac{\partial F}{\partial \phi}\left(1 + Br^{2k} + \frac{\Omega^*}{r^2}\right) + \frac{\partial F}{\partial r}\left(Ar^{2k+1} + Lr^{4k+1} + \frac{R^*}{r}\right) = AF^{2k+1} + LF^{4k+1},$$

where $A, B, L, R^*(r, \phi), \Omega^*(r, \phi)$ *have the same significance as in* (8.2), *has a solution*

$$(9.31) \qquad F = r + \sum F_{ij}(r) \cos^i\phi \sin^j\phi$$

analytic in r *for* $r > 0$, *of class* C^∞ *for* $r \geqq 0$ *and such that*

$$\lim_{r=0} \frac{F_{ij}}{r^p} = 0$$

for any positive integer p.

Consider the two cases $L = 0$ and $L \neq 0$.

Case 1. $L = 0$.

If $L = 0$, the partial differential equation (9.3) becomes

$$(9.4) \qquad \frac{\partial F}{\partial \phi}\left(1 + Br^{2k} + \frac{\Omega^*}{r^2}\right) + \frac{\partial F}{\partial r}\left(Ar^{2k+1} + \frac{R^*}{r}\right) = AF^{2k+1}.$$

Make the substitution

$$(9.41) \qquad f = \frac{e^{\frac{-1}{F^{2k}}}}{e^{-\frac{1}{r^{2k}}}}$$

in the partial differential equation (9.4). We get

$$(9.42) \qquad \frac{\partial f}{\partial \phi}\left(1 + Br^{2k} + \frac{\Omega^*}{r^2}\right) + \frac{\partial f}{\partial r}\left(Ar^{2k+1} + \frac{R^*}{r}\right) + \frac{2kR^*}{r^{2k+2}} f = 0$$

and $f = 1 + \cdots$, where the part of the expression for f which is omitted vanishes with r to a power higher than the first.

Now introduce the variables

$$(9.43) \qquad \begin{aligned} \psi &= \phi + \frac{1}{2kAr^{2k}} - \frac{B}{A}\log r, \\ \rho &= r. \end{aligned}$$

The partial differential equation (9.4) becomes

$$\frac{\partial f}{\partial \psi}\left(\frac{\bar{\Omega}^*}{\rho^2} - \frac{\bar{R}^*(1 + B\rho^{2k})}{A\rho^{2k+2}}\right) + \frac{\partial f}{\partial \rho}\left(A\rho^{2k+1} + \frac{R^*}{\rho}\right) + \frac{2k\bar{R}^*}{\rho^{2k+2}} f = 0. \tag{9.44}$$

The expressions $R^*(r, \phi)$, $\Omega^*(r, \phi)$ above become, in terms of ρ, ψ, $\bar{R}^*(\rho, \psi)$, $\bar{\Omega}^*(\rho, \psi)$. These expressions have the same properties in terms of ρ and ψ that the original expressions, R^*, Ω^*, have in terms of r and ϕ.

Finally write (9.44) as

$$\frac{\partial f}{\partial \rho} = \frac{-2k\bar{R}^*}{\rho^{2k+2}\left(A\rho^{2k+1} + \dfrac{\bar{R}^*}{\rho}\right)} f + \frac{\dfrac{(1 + B\rho^{2k})\bar{R}^*}{A\rho^{2k+1}} - \dfrac{\bar{\Omega}^*}{\rho^2}}{A\rho^{2k+1} + \dfrac{\bar{R}^*}{\rho}} \frac{\partial f}{\partial \psi}, \tag{9.45}$$

where $f(0, \psi) = 1$.

Now let us consider the other case.

Case 2. $L \neq 0$.

Here make in (9.3) the substitution

$$A \log u = -\frac{1}{F^{2k}} + \frac{L}{A} \log\left[\frac{1}{F^{2k}} + \frac{L}{A}\right] + \frac{1}{r^{2k}} - \frac{L}{A} \log\left[\frac{1}{r^{2k}} + \frac{L}{A}\right]. \tag{9.5}$$

The partial differential equation becomes

$$\frac{\partial u}{\partial \phi}\left(1 + Br^{2k} + \frac{\Omega^*}{r^2}\right) + \frac{\partial u}{\partial r}\left(Ar^{2k+1} + Lr^{4k+1} + \frac{R^*}{r}\right) = -\frac{2k\dfrac{R^*}{r}}{Ar^{2k+1} + Lr^{4k+1}}, \tag{9.51}$$

where $u = 1 + \cdots$. Introduce in (9.51) the new variables

$$\psi = \phi + \left\{\frac{AB - L}{A} \log \frac{A + Lr^{2k}}{r^{2k}} + \frac{1}{r^{2k}}\right\} \frac{1}{2kA}, \tag{9.52}$$

$$\rho = r,$$

whence

$$\frac{\partial u}{\partial \rho} = \frac{1}{A\rho^{2k+1} + L\rho^{4k+1} + \dfrac{\bar{R}^*}{\rho}}\left\{\frac{2k\dfrac{\bar{R}^*}{\rho}u}{A\rho^{2k+1} + L\rho^{4k+1}} + \frac{\partial u}{\partial \psi}\left[\frac{\bar{\Omega}^*}{\rho^2} - \frac{1 + B\rho^{2k}}{A\rho^{2k+1} + L\rho^{4k+1}}\frac{R^*}{\rho}\right]\right\}, \tag{9.53}$$

where $\bar{\Omega}^*(\rho, \psi)$, $\bar{R}^*(\rho, \psi)$ arise from $\Omega^*(r, \phi)$, $R^*(r, \phi)$.

By virtue of the lemma below, each of the partial differential equations (9.45) and (9.53) has a solution in terms of $\cos \psi$, $\sin \psi$, ρ of the form (9.71) below, convergent for

$$\rho \leqq \rho_0, \qquad 0 \leqq \psi \leqq 2\pi,$$

where the coefficients of the $\cos^i\psi \sin^j\psi$ have the property (9.72) and where $u(0, \psi) = 1$.

Returning to the variables r, ϕ by means of (9.43) or (9.52), the corresponding function $f(r, \phi)$ is also of the form (9.71) convergent for $r \leqq r_0$, $0 \leqq \phi \leqq 2\pi$.

Finally introduce the function $F(r, \phi)$ by means of (9.5) or (9.4). We find in both cases that this function, a solution of (9.3), likewise can be expressed in the form (9.71) with the above properties. Consequently we have the lemma established.

If the same transformation (9.43) or (9.52) is made in the partial differential equation (9.6) the following lemma can be established:

LEMMA. *The partial differential equation*

$$(9.6) \qquad \frac{\partial G}{\partial \phi}\left(1 + Br^{2k} + \frac{\Omega^*}{r^2}\right) + \frac{\partial G}{\partial r}\left(Ar^{2k+1} + Lr^{4k+1} + \frac{R^*}{r}\right) = 1 + BF^{2k}$$

has a solution

$$G = \phi + \sum g_{ij} \cos^i\phi \sin^j\phi$$

analytic in r for $r > 0$, of class C^∞ in r for $r \geqq 0$ and such that

$$\lim_{r=0} \frac{g_{ij}}{r^p} = 0,$$

p any integer.

Now writing $w = f_1 = F \cos G$, $z = g_1 = F \sin G$, and expressing F and G in terms of u and v by $u = r \cos \phi$, $v = r \sin \phi$, the theorem of this section can be established.

Finally, the lemma appealed to above can be stated as follows:

LEMMA. *In the partial differential equation*

$$(9.7) \qquad \frac{\partial f}{\partial \rho} = S(\rho, \theta) f + T(\rho, \theta) \frac{\partial f}{\partial \theta},$$

let the functions $S(\rho, \theta)$, $T(\rho, \theta)$ be analytic in ρ and θ, $\rho > 0$ class C^∞, $\rho \geqq 0$ expressible as

$$(9.71) \qquad \begin{aligned} &\text{(a)} \quad S(\rho, \theta) = \sum S_i(\rho)\theta^i, \\ &\text{(b)} \quad S(\rho, \theta) = \sum S_{ij}(\rho) \cos^i\theta \sin^j\theta \end{aligned}$$

both uniformly convergent for $\rho_0 \geqq \rho \geqq 0$ and such that

$$(9.72) \qquad \lim_{\rho=0} \frac{S_i(\rho)}{\rho^p} = 0, \qquad \lim_{\rho=0} \frac{S_{ij}(\rho)}{\rho^p} = 0$$

for any integer p, with similar properties for $T(\rho, \theta)$. Then there exists a function $f(\rho, \theta)$, analytic in ρ and θ, $\rho > 0$, of class C^∞, $\rho \geqq 0$, expressible as $S(\rho, \theta)$ is above,

(9.71) *and* (9.72), *and satisfying the partial differential equation* (9.7) *and* $f(0, \theta) = 1$.

This lemma can be established by setting up the sequence

$$\begin{aligned} f_1 &= 1, \\ f_{i+1} &= 1 + \int_0^\rho \left\{ S(\rho, \theta) f_i + T(\rho, \theta) \frac{\partial f_i}{\partial \theta} \right\} d\rho. \end{aligned} \tag{9.73}$$

The functions S and T are expressed as under (a) above to prove the convergence. The proof follows in the way used for ordinary differential equations.[9]

10. **Conclusion.** Combining the transformations of §8 and §9, we take the system of differential equations (6.1) into the final form (9.2) using a transformation analytic in the open region, $x^2 + y^2 > 0$, and of class C^∞ in the closed region $x^2 + y^2 \geqq 0$ for $(x^2 + y^2)$ not greater than, say, d^2.

As indicated above, the form (6.1) is attained by an analytic transformation from the initial form (1.1) or by an analytic transformation and the change of parameter indicated in §4.

Consequently the final form of the system of differential equations (9.2) is attained from the initial system (1.1) by a transformation analytic in the open region, of class C^∞ in the closed region including the origin, or by such a transformation and a change of parameter.

UNIVERSITY OF CINCINNATI.

[9] Picard, *Traité d'Analyse*, vol. 2, 1925, p. 368.

THE MINIMA OF FUNCTIONALS WITH ASSOCIATED SIDE CONDITIONS

By Herman H. Goldstine

In his doctoral dissertation the author obtained a generalization of the calculus of variations which does not include either the problem of Lagrange with fixed end-points or the more general problem of Bolza. That is to say, the theory therein presented includes only the ordinary problem of the calculus of variations and certain non-calculus of variations problems which have fixed end-points and no side conditions.[1] To remedy this defect a more general situation is considered in the present paper; more specifically it is proposed to find conditions that a functional having certain differentiability properties be a minimum in a class of functions satisfying a system of integro-differential equations and passing through two fixed points.

This problem is transformed into one having only generalized end conditions and is then treated by a method which is a generalization of the technique adopted in the author's thesis. Analogues of the Lagrange multiplier rule, the Clebsch condition, and the Jacobi-Mayer condition are obtained for the transformed problem.

The analogue of the Jacobi-Mayer condition is especially interesting, since the theory of the fixed end-point problem of Lagrange with integro-differential side conditions does not contain such a condition.[2] Moreover this condition does not reduce to the ordinary Jacobi condition for the simple problem of the calculus of variations, as has been shown.[3]

1. **The problems and the transformation.** We shall start with the following problem: to find necessary conditions that an arc

$$\xi_i = \xi_i(s) \qquad (i = 1, \cdots, n; 0 \leqq s \leqq 1)$$

minimize a functional I in the class of arcs satisfying the integro-differential conditions

Received October 26, 1936.

[1] *Conditions for a minimum of a functional*, Chicago Doctoral Dissertation (1936); it is expected that this paper will soon appear in the third volume of the *Contributions to the Calculus of Variations* (Chicago). In the subsequent footnotes it will be indicated as Paper I.

[2] See, e.g., L. M. Graves, *A transformation of the problem of Lagrange in the calculus of variations*, Transactions of the American Mathematical Society, vol. 35 (1933), pp. 675–682.

[3] See Paper I, pp. 32–37.

$$\varphi_\alpha[x, \xi(x), u(x)] = 0 \qquad (\alpha = 1, \cdots, m < n),$$

(1)

$$u_\gamma(x) = \int_0^x P_\gamma[x, s, \xi(s)]\, ds \qquad (\gamma = n + 1, \cdots, n + q),$$

(2)

$$u_i(x) = a_i + \int_0^x \xi_i(s)\, ds \qquad (i = 1, \cdots, n)$$

and the end conditions $u_i(1) = b_i$ $(i = 1, \cdots, n)$. The functions φ_α are assumed to have continuous partial derivatives of the second order with respect to all their arguments in a given region $\mathfrak{R}_1$ of $(2n + q + 1)$-dimensional space. The functions P_γ are supposed to possess the same differentiability properties as the functions φ_α on a region $\mathfrak{R}_2$ of $(n + 2)$-dimensional space. The curves to be discussed belong to a certain region $\mathfrak{C}_{0n}$ of the space $\mathfrak{C}_n$ of n-uples of continuous functions defined on $0 \leqq x \leqq 1$. Moreover on this region $\mathfrak{C}_{0n}$ the functional I has a second differential.[4]

The region $\mathfrak{R}$ of $(2n + q + 2)$-dimensional space is, by definition, the set of all points $(x, s, y_1', \cdots, y_n', u_1, \cdots, u_{n+q})$ for which the point $(x, s, y_1', \cdots, y_n')$ is in $\mathfrak{R}_2$ and the point $(x, y_1', \cdots, y_n', u_1, \cdots, u_{n+q})$ is in $\mathfrak{R}_1$. Then the region $\mathfrak{C}_{0n}$ has the property that the point

$$[x, s, \xi(s), u(s)] \qquad (0 \leqq s \leqq x \leqq 1)$$

is interior to the region $\mathfrak{R}$ whenever the element $(\xi_1, \cdots, \xi_n)$ is in $\mathfrak{C}_{0n}$. Lastly along the minimizing "curve"

$$\xi_{0i} = \xi_{0i}(s), \qquad u_\delta = u_\delta(s) \qquad (i = 1, \cdots, n; \delta = 1, \cdots, n + q),$$

the matrix of partial derivatives $(\varphi_{\alpha y_i'})$ $(\alpha = 1, \cdots, m;\ i = 1, \cdots, n)$ has rank m.

The second problem to be treated is the following: to find necessary conditions that an element $\xi_0 = (\xi_{01}, \cdots, \xi_{0r})$, lying in a certain region $\mathfrak{C}_{0r}$ of the space of r-uples of continuous functions defined on $0 \leqq x \leqq 1$ minimize a functional J in a class of elements satisfying functional conditions

$$A_\mu(\xi_0) = 0 \qquad (\mu = 1, \cdots, p).$$

The functionals J and A_μ are supposed to have second differentials at each point of $\mathfrak{C}_{0r}$.

It is simple to verify that the transformation adopted by Graves in treating the problem of Lagrange[5] carries the first problem into a special case of the second; moreover the transformation is such that necessary conditions for one problem are necessary for the other.

[4] For definitions of the terms used above see, e.g., L. M. Graves, *Topics in the functional calculus*, Bulletin of the American Mathematical Society, vol. 41 (1935), pp. 641–662; T. H. Hildebrandt and L. M. Graves, *Implicit functions and their differentials in general analysis*, vol. 29 (1927), pp. 127–153; or Paper I.

[5] *A transformation of the problem of Lagrange in the calculus of variations*, loc. cit.

However, before proceeding to a treatment of the latter problem, we shall show that it can be transformed into an equivalent problem which is notationally simpler to analyse. To effect this transformation we make the following definitions:

$$\xi(x) = \xi_i(x - i + 1) \qquad (i - 1 \leqq x < i;\, i = 1, \cdots, r), \tag{3}$$
$$\xi(r) = \xi_r(1),$$

$$K(\xi) = J[(\xi_1, \cdots, \xi_r)], \quad B_\mu(\xi) = A_\mu[(\xi_1, \cdots, \xi_r)] \qquad (\mu = 1, \cdots, p), \tag{4}$$

where ξ is given by equations (3). Then the domain of K and B_μ is a region $\mathfrak{C}_0$ of the space of functions ξ which are defined on $0 \leqq x \leqq r$ and are continuous except possibly at $x = i$ $(i = 1, \cdots, r - 1)$, where they have finite right and left hand limits; we define the norm of an element ξ to be the largest of the norms of its *sections*,[6] whence K and B_μ have second differentials on this region.

Making use of the Riesz theorem[7] and the differentiability of K, one can verify that there exist *regular* functions κ and β_μ of limited variation[8] such that $\kappa(0) = \beta_\mu(0) = 0$ $(\mu = 1, \cdots, p)$, κ and β_μ are continuous at $x = i$ $(i = 1, \cdots, r - 1)$, and

$$dK(\xi_0; \zeta) = \int_0^r \zeta(x)\, d\kappa(x), \qquad dB_\mu(\xi_0; \zeta) = \int_0^r \zeta(x)\, d\beta_\mu(x) \qquad (\mu = 1, \cdots, p) \tag{5}$$

for every *admissible variation* ζ, i.e., for every function ζ whose sections are continuous functions.

2. **The multiplier rule.**[9] Since K is a minimum at ξ_0 in the class of all ξ in $\mathfrak{C}_0$ satisfying the equations $B_\mu(\xi) = 0$, the usual considerations of the calculus of variations suffice to show that there exist constants $l_0, c_1, \cdots, c_p$, not all zero, such that

$$l_0\, dK(\xi_0; \zeta) + c_\mu\, dB_\mu(\xi_0; \zeta) = 0$$

for every admissible variation ζ. Hence defining $\lambda(x)$ to be $l_0\kappa(x) + c_\mu\beta_\mu(x)$, we have

$$\int_0^r \zeta(x)\, d\lambda(x) = 0, \tag{6}$$

from which follows

[6] By definition, the i-th section ξ_i of ξ is $\xi(x)$, where $i - 1 \leqq x < i$, and $\xi_i(i) = \lim_{x<i} \xi(x)$; hence each section of ξ is a continuous function on a finite and closed interval.

[7] See, e.g., S. Banach, *Théorie des Opérations Linéaires*, Warsaw, 1932, pp. 59 ff.

[8] A function f of limited variation will be said to be regular in case

$$f(x) = [f(x - 0) + f(x + 0)]/2;$$

every function of limited variation can be so regularized.

[9] See, e.g., H. Hahn, *Ueber die Lagrange'sche Multiplikatorenmethode*, Sitzungsberichte der Akademie Wien, vol. 131 (1922), pp. 531–550.

THE MULTIPLIER RULE. *If the function $\xi_0(x)$ minimizes the functional K in the class of all functions ξ in $\mathfrak{C}_0$ which satisfy the equations*

$$B_\mu(\xi) = 0,$$

then there exist constants $l_0, c_1, \cdots, c_p$, not all zero, such that

$$\lambda(x) = l_0\kappa(x) + c_\mu\beta_\mu(x) = 0 \qquad (0 \leqq x \leqq r), \tag{7}$$

where κ and β_μ are the functions appearing in equations (5).

To prove the theorem it suffices to note that equation (6) implies the vanishing of λ at $x = r$ and to use the method indicated in the author's thesis.[10]

An element ξ in $\mathfrak{C}_0$ will be said to be *normal* in case there exist p admissible variations ζ_σ $(\sigma = 1, \cdots, p)$ such that the determinant $| dB_\mu(\xi; \zeta_\sigma) |$ $(\mu, \sigma = 1, \cdots, p)$ does not vanish. By a proof similar to the one made in the calculus of variations it can be shown that an element ξ is normal if and only if it has no set of multipliers $l_0 = 0, c_1, \cdots, c_p$, with which it satisfies equation (7). As usual we can take $l_0 = 1$ for a normal element and hence uniquely determine the remaining multipliers.

3. The second differential.

The usual considerations show that, if our minimizing element ξ_0 is normal, then $dK(\xi_0; \zeta) = 0$ and $d^2K(\xi_0; \zeta, \zeta) \geqq 0$ for every admissible variation satisfying

$$dB_\mu(\xi_0; \zeta) = 0 \qquad (\mu = 1, \cdots, p). \tag{8}$$

We find it convenient to impose a restriction on the functionals B_μ.

(H.1) *The matrix $(dB_\mu(\xi_0; x^{\sigma-1}))$ $(\mu, \sigma = 1, \cdots, p)$ is of rank p.*

Therefore the vectors

$$dB_\mu(\xi_0; 1), dB_\mu(\xi_0; x), \cdots, dB_\mu(\xi_0; x^{p-1}) \qquad (\mu = 1, \cdots, p) \tag{9}$$

are linearly independent, and ξ_0 is a normal element. Then there is no loss of generality in supposing that

$$\beta_\mu(0) = 0, \qquad dB_\mu(\xi_0; x^{\sigma-1}) = \delta_{\mu\sigma} \qquad (\mu, \sigma = 1, \cdots, p), \tag{10}$$

where $(\delta_{\mu\sigma})$ is the Kronecker delta. Since the vectors (9) are linearly independent, there exists a non-singular matrix $(c_{\mu\sigma})$ such that

$$\delta_{\mu\sigma} = c_{\mu j}\, dB_j(\xi_0; x^{\sigma-1}).$$

Then the functionals $C_\mu = c_{\mu j}B_j$ have the desired properties; the system of equations $dC_\mu(\xi_0; \zeta) = 0$ has the same solutions as the system $dB_\mu(\xi_0; \zeta) = 0$, and $dC_\mu(\xi_0; x^{\sigma-1}) = \delta_{\mu\sigma}$. It will be supposed throughout the sequel that the system (8) has been put into this canonical form. Next it will be shown that

LEMMA 1. *If $d^2K(\xi_0; \nu, \nu) \geqq 0$ for every function ν whose sections have a continuous p-th derivative and which satisfies $dB_\mu(\xi_0; \nu) = 0$, then $d^2K(\xi_0; \zeta, \zeta) \geqq 0$ for every admissible variation ζ satisfying $dB_\mu(\xi_0; \zeta) = 0$.*

[10] Paper I, pp. 21 ff.

Suppose that there is an admissible variation ζ_0 for which the lemma is false. Then applying the Weierstrass approximation theorem to the sections of ζ_0, we get a sequence of functions $\{\nu_n\}$ whose sections have continuous p-th derivatives and which converges in the sense of the norm to ζ_0. Let $\nu_{1n}(x) = \nu_n(x) + c_{1n} + c_{2n}x + \cdots + c_{pn}x^{p-1}$. Then for each n and each set $c_{1n}, \cdots, c_{pn}$ of constants, ν_{1n} has the desired differentiability properties. Furthermore by choosing the $c_{1n}, \cdots, c_{pn}$ so that $c_{\mu n} = -dB_\mu(\xi_0; \nu_n)/dB_\mu(\xi_0; x^{\mu-1})$, we see that ν_{1n} satisfies $dB_\mu(\xi_0; \nu) = 0$ for every n. Moreover, since ζ_0 satisfies these equations, it is evident that for each μ the sequence $c_{\mu n}$ converges to zero; ν_{1n} converges in the sense of the norm to ζ_0. This leads to a contradiction.[11]

We proceed to a determination of the second differential. It can be seen from the proofs of Fréchet[12] that there exists a symmetric function $k(x, y)$ of limited variation in the sense of Fréchet[13] such that $k(x, 0) = 0 = k(0, y)$ $(0 \leqq x \leqq r; 0 \leqq y \leqq r)$, k is continuous in both variables together along the lines $x = i, y = j$ $(i, j = 1, \cdots, r-1)$, and for every admissible ζ_1 and ζ_2:[14]

$$d^2K(\xi_0; \zeta_1, \zeta_2) = \int_0^r \int_0^r \zeta_1(x)\zeta_2(y)d_x d_y k(x, y), \tag{11}$$

Then adopting the method of the author's thesis,[15] we define functions k_μ:

$$k_1(x, y) = k(x, y) + k(r, r)\frac{\beta_1(x)\beta_1(y)}{\beta_1(r)\beta_1(r)} - k(x, r)\frac{\beta_1(y)}{\beta_1(r)} - k(r, y)\frac{\beta_1(x)}{\beta_1(r)},$$

$$\begin{aligned} k_j(x, y) = {} & \int_0^x ds \int_0^y k_{j-1}(s, t)dt \\ & + \int_0^r ds \int_0^r k_{j-1}(s, t)\,dt \frac{\int_0^x (x-s)^{j-2}\beta_j(s)ds \int_0^y (y-s)^{j-2}\beta_j(s)ds}{\int_0^r (r-s)^{j-2}\beta_j(s)ds \int_0^r (r-s)^{j-2}\beta_j(s)ds} \\ & - \int_0^x ds \int_0^r k_{j-1}(s, t)dt \frac{\int_0^y (y-s)^{j-2}\beta_j(s)ds}{\int_0^r (r-s)^{j-2}\beta_j(s)ds} \\ & - \int_0^r ds \int_0^y k_{j-1}(s, t)dt \frac{\int_0^x (x-s)^{j-2}\beta_j(s)ds}{\int_0^r (r-s)^{j-2}\beta_j(s)ds}. \end{aligned}$$

[11] Paper I, pp. 11 ff.

[12] *Sur les fonctionelles bilinéaires*, Transactions of the American Mathematical Society, vol. 16 (1915), pp. 215–244.

[13] Ibid., pp. 223 ff.

[14] See, e.g., Paper I, pp. 4 ff. The continuity property stated above results from the arbitrariness with which the values of k may be chosen.

[15] See Paper I, pp. 6–8.

Then the functions k_μ ($\mu = 1, \cdots, p$) are symmetric, vanish on the boundary of the square ($0 \leqq x \leqq r$; $0 \leqq y \leqq r$); k_1 is bounded, Riemann integrable,[16] and continuous on the lines $x = i$, $y = j$ ($i, j = 1, \cdots, r-1$). Moreover by successively integrating by parts,[17] and using equations (9) and (10), we find that

$$(12) \qquad d^2K(\xi_0; \nu_1, \nu_2) = \int_0^r \int_0^r \nu_1^{(p)}(x) k_p(x, y) \nu_2^{(p)}(y)\,dx\,dy \qquad (\nu_1 \text{ in } \mathfrak{N}, \nu_2 \text{ in } \mathfrak{N}),$$

where $\mathfrak{N}$ is the class of all admissible variations ν satisfying equations (8), such that the sections of ν have continuous p-th derivatives.

We find it convenient to make a second restriction on the generality of our problem.

(H.2) *If $p = 1$, it is assumed that k_1 is continuous in both variables together.*

4. **The Jacobi-Mayer condition.** We proceed to establish the

Analogue of the Jacobi-Mayer Condition. *If ξ_0 is a normal minimizing element for K, then the integral equation*

$$(13) \qquad \int_0^r k_p(x, y)\varphi(y)\,dy = \sigma\varphi(x),$$

where k_p is the unique continuous function appearing in equation (12), *can have no negative characteristic values. Moreover if $d^2K(\xi_0; \zeta, \zeta)$ is not equal to zero for every admissible variation satisfying $dB_\mu(\xi_0; \zeta) = 0$ ($\mu = 1, \cdots, p$), then equation* (13) *always has at least one non-zero characteristic value.*[18]

We prove first that k_p is unique. Let κ_1 and κ_2 be two continuous functions effective in equation (12) and let $\kappa = \kappa_1 - \kappa_2$. Then

$$\int_0^r \int_0^r \nu_1^{(p)}(x)\kappa(x, y)\nu_2^{(p)}(y)\,dx\,dy = 0 \qquad (\nu_1 \text{ in } \mathfrak{N}; \nu_2 \text{ in } \mathfrak{N}).$$

Consider any two continuous functions

$$\varphi_1(x), \qquad \varphi_2(x);$$

then define two functions $\nu_1(x)$, $\nu_2(x)$ as follows:

$$(14) \quad \nu_i(x) = \int_0^x \int_0^{x_1} \cdots \int_0^{x_{p-1}} \varphi_i(x_p)\,dx_p \cdots dx_1 + c_{p-1,i}x^{p-1} + \cdots + c_{1,i}x + c_{0,i}$$
$$(i = 1, 2; 0 \leqq x \leqq r).$$

[16] Ibid., pp. 8–10.

[17] Ibid., p. 11.

[18] Ibid., pp. 23–24, 26–41. In this reference the condition above is obtained for the case $r = p = 1$, $\beta_1(x) = x$; the result is interpreted for two special cases, one of which is the fixed end-point problem of the calculus of variations.

We can determine the numbers $c_{0,i}, \cdots, c_{p-1,i}$ so that they satisfy the equations

$$0 = \int_0^r \int_0^x \int_0^{x_1} \cdots \int_0^{x_{p-1}} \varphi_i(x_p)dx_p \cdots dx_1 d\beta_\mu(x) + c_{\mu-1,i}dB_\mu(\xi_0; x^{\mu-1})$$
$$(i = 1, 2; \mu = 1, \cdots, p).$$

Then both ν_1 and ν_2 are in $\mathfrak{N}$. Hence we have proved that

$$\int_0^r \int_0^r \varphi_1(x)\kappa(x, y)\varphi_2(y)dxdy = 0$$

for every pair (φ_1, φ_2) of continuous functions; this implies that κ is identically zero.

Suppose next that σ is a negative characteristic value for equation (13). Then in (14) replace φ_i by φ, a normed characteristic solution corresponding to σ. Hence we have defined a function ν in $\mathfrak{N}$ such that

$$\int_0^r k_p(x, y)\nu^{(p)}(y)dy = \sigma\nu^{(p)}(y), \qquad \int_0^r [\nu^{(p)}(x)]^2\, dx = 1.$$

This leads immediately to a contradiction.

Finally if k_p is identically zero, $d^2K(\xi_0; \nu, \nu) = 0$ for every ν in $\mathfrak{N}$, whence by a method analogous to that employed in Lemma 1 we find that

$$d^2K(\xi_0; \zeta, \zeta) = 0$$

for every admissible variation satisfying equations (8). Consequently the last statement in the condition follows immediately from the symmetry of k_p and the Hilbert-Schmidt theory.

COROLLARY 1. *If ξ_0 is a normal minimizing element for K, then*

$$k_{p,n}(x, x) \geqq 0 \qquad (n = 1, 2, \cdots; 0 \leqq x \leqq r),$$

where $k_{p,n}(x, y)$ is the n-th iterated kernel for $k_p(x, y)$.

The proof is the same as the one in the author's thesis.[19] Apparently neither the condition above nor its corollary appear in the literature for any special case of the theory herein developed.[20]

5. The Clebsch condition.

We are able to state a simple generalization of the familiar Clebsch necessary condition.

ANALOGUE OF THE CLEBSCH CONDITION. *Let ξ_0 be a normal element which minimizes K in the class of all admissible curves satisfying $B_\mu(\xi) = 0$, and let*

$$\eta(x) = \psi(x \mid \zeta) \qquad (0 \leqq x \leqq r) \tag{15}$$

be a Lebesgue square integrable function of x for each admissible variation ζ satis-

[19] See Paper I, p. 24.

[20] I.e., excepting the case $r = p = 1$, $\beta_1(x) = x$, which appears in the author's dissertation. See, e.g., Graves, *A transformation of the problem of Lagrange in the calculus of variations*, loc. cit., p. 675.

fying $dB_\mu(\xi_0 ; \zeta) = 0$ $(\mu = 1, \cdots, p)$; *further suppose that* ψ *is a homogeneous functional of* ζ *for* $0 \leqq x \leqq r$. *Then the lower bound* $B(\xi_0)$ *of* $d^2K(\xi_0 ; \zeta, \zeta)$ *in the class of all admissible variations* ζ *such that* $dB_\mu(\xi_0 ; \zeta) = 0$ $(\mu = 1, \cdots, p)$ *and*

$$\int_0^r \eta^2(x)\,dx = 1 \tag{16}$$

is finite; this number is also the lower bound of $d^2K(\xi_0 ; \nu, \nu)$ *for all* ν *in* $\mathfrak{N}$ *satisfying equation* (16).

The first part of this theorem follows immediately from the homogeneity of ψ as a function of ζ and the normality of ξ_0. In the proof of Lemma 1 a method for constructing sequences of elements in $\mathfrak{N}$ which converge to any admissible ζ with $dB_\mu(\xi_0 ; \zeta) = 0$ was given. By means of this convergence property an indirect proof of the latter part of the condition can be made readily.[21]

In Paper I the author has shown how the finiteness of the bound $B(\xi_0)$ is equivalent to the Legendre condition in the special case of the ordinary problem of the calculus of variations. Further, the condition above in conjunction with the Jacobi-Mayer condition implies that the bound $B(\xi_0)$ is both finite and non-negative. The problem of obtaining sufficient conditions has been solved only for the restricted case which is considered in the author's thesis.

The University of Chicago.

[21] See Paper I, p. 23.

ON THE ZEROS OF JACOBI POLYNOMIALS, WITH APPLICATIONS

By M. S. Webster

Introduction. This paper deals principally with the generalized Jacobi polynomials

$$J_n(x;\alpha,\beta) \equiv J_n(x) \equiv (1+x)^{1-\alpha}(1-x)^{1-\beta}\frac{d^n}{dx^n}[(1+x)^{n+\alpha-1}(1-x)^{n+\beta-1}]$$

$$\equiv (-1)^n n!\binom{2n+\alpha+\beta-2}{n}\phi_n(x;\alpha,\beta); \tag{1}$$

$$\phi_n(x;\alpha,\beta) \equiv \phi_n(x) \equiv x^n - S_n x^{n-1} + \cdots \qquad (n = 0, 1, 2, \cdots),$$

defined (except for constant factors) for all real α, β as the polynomial solutions of the differential equation

$$\begin{aligned}(1-x^2)J_n''(x) + [\alpha - \beta - (\alpha+\beta)x]J_n'(x)&\\ + n(n+\alpha+\beta-1)J_n(x) = 0 &\qquad (n = 0, 1, \cdots).\end{aligned} \tag{2}$$

For arbitrary α, β several authors have discussed the number of real zeros of $J_n(x;\alpha,\beta)$. Stieltjes [1][1] gave a method of finding the number of zeros in the intervals $(-\infty, -1)$, $(-1, 1)$, $(1, \infty)$ but he stated the result only when $\alpha, \beta > 0$. Shibata [2] gave a table for the number of zeros when they are all real and α, β are not negative integers or zero. Lawton [3] gave complete results for the closed interval $(-1, 1)$ when n is sufficiently large. The results of Hilbert [4], Klein [5], Van Vleck [6], and Hurwitz [7] for the zeros of the hypergeometric function may also be applied to Jacobi polynomials.

Here we find the number of zeros of $J_n(x;\alpha,\beta)$ inside the intervals $(-\infty, -1)$, $(-1, 1)$, $(1, \infty)$ and, in addition, at $x = \pm 1$ (from which the number of imaginary zeros is easily obtained) when α, β are arbitrary. The method employed is new and other properties of $J_n(x;\alpha,\beta)$ are developed as well.

In case $\alpha, \beta > 0$, the $J_n(x)$ form, as is known, an orthogonal system

$$\int_{-1}^{1} p(x)J_m(x)J_n(x)dx = 0, \qquad p(x) = (1+x)^{\alpha-1}(1-x)^{\beta-1} \tag{3}$$

$$(m \neq n;\ m, n = 0, 1, \cdots).$$

Received December 7, 1936; presented to the American Mathematical Society, April 20, 1935. The author wishes to express his gratitude to Professor J. Shohat for his valuable suggestions.

[1] The numbers in brackets refer to the bibliography at the end of the paper.

Here the zeros $x_{i,n}(\alpha, \beta) \equiv x_{i,n}$ ($i = 1, 2, \cdots, n$) of $J_n(x)$ have the following properties:

$$x_{1,n} \to -1, \qquad x_{n,n} \to 1 \ (n \to \infty);$$

$$-1 < x_{1,n+1} < x_{1,n} < x_{2,n+1} < \cdots < x_{n,n} < x_{n+1,n+1} < 1.$$

We obtain upper and lower bounds for $x_{i,n}$ by the geometrical method which E. R. Neumann [8], Winston [9], and W. Hahn [10] used for Laguerre and Hermite polynomials. We further employ a superior method, based on Markoff's theorem [11], which enables us to find better bounds and even asymptotic expressions ($n \to \infty$) for certain $x_{i,n}$. We then apply these results to the coefficients $H_{i,n}$ in the mechanical quadratures formula

$$(4) \quad \int_{-1}^{1} p(x)f(x)dx = \sum_{i=1}^{n} H_{i,n}f(x_{i,n}) + R_n(f), \qquad H_{i,n} = \int_{-1}^{1} \frac{p(x)\phi_n(x)dx}{(x - x_{i,n})\phi_n'(x_{i,n})},$$

with the characteristic property that

$$\int_{-1}^{1} p(x)G_{2n-1}(x)dx = \sum_{i=1}^{n} H_{i,n}G_{2n-1}(x_{i,n}), \qquad \text{i.e.,} \quad R_n(G_{2n-1}(x)) \equiv 0,$$

$G_s(x)$ denoting an arbitrary polynomial of degree $\leqq s$.

We consider also the coefficients $H_{i,n}$ for Laguerre polynomials.

Finally, we establish the theorem, recently proved in a different way by Meixner [12], that the system of Hermite polynomials is the only case where systems of orthogonal and Appell [13] polynomials coincide.

1. The following relations, being identities in α, β, remain valid for the generalized $J_n(x)$ (as well as for the orthogonal case), $n = 1, 2, \cdots$.

$$(5) \quad J_n'(x; \alpha, \beta) = -n(n + \alpha + \beta - 1)J_{n-1}(x; \alpha + 1, \beta + 1)$$

$$(6) \quad \begin{aligned} J_n(-x; \alpha, \beta) &\equiv (-1)^n J_n(x; \beta, \alpha); \\ x_{i,n}(\alpha, \beta) &= -x_{n-i+1,n}(\beta, \alpha) \qquad (i = 1, 2, \cdots, n) \end{aligned}$$

$$(7) \quad J_n(-1; \alpha, \beta) = 2^n \cdot n! \binom{n + \alpha - 1}{n}; \; J_n(1; \alpha, \beta) = (-2)^n \cdot n! \binom{n + \beta - 1}{n}.$$

With Lawton, we let $\psi(x) \equiv (1 + x)^{n+\alpha-1}(1 - x)^{n+\beta-1}$, so that

$$\frac{d^n}{dx^n} [(1 + x)\psi(x)] \equiv (1 + x) \frac{d^n}{dx^n} \psi(x) + n \frac{d^{n-1}}{dx^{n-1}} \psi(x).$$

In virtue of (1, 2), we obtain by successive differentiation

$$(8) \quad [n + \alpha + \beta - 1][J_n(x; \alpha + 1, \beta) - J_n(x; \alpha, \beta)] = (x - 1)J_n'(x; \alpha, \beta),$$

$$(9) \quad [n(x + 1) + 2\alpha]J_n(x; \alpha + 1, \beta) - 2(n + \alpha) J_n(x; \alpha, \beta) = (x^2 - 1)J_n'(x; \alpha + 1, \beta).$$

By repeated use of (5, 7), we have

$$
(10)\qquad
\begin{aligned}
J_n(x;\alpha,\beta) &\equiv \sum_{i=0}^{n} J_n^{(i)}(-1;\alpha,\beta)\,\frac{(x+1)^i}{i!} \equiv \sum_{i=0}^{n}\gamma_i(x+1)^i\\
&\equiv n!\sum_{i=0}^{n}(-1)^i\,2^{n-i}\binom{n+i+\alpha+\beta-2}{i}\binom{n+\alpha-1}{n-i}(x+1)^i.
\end{aligned}
$$

Hereafter, r represents a negative integer or zero; $[x] = 0$ if $x < 1$, and equals the largest positive integer $\leqq x$ if $x \geqq 1$. Introduce non-negative integers p, q, n_0 as follows:

$$
p = [1-\alpha],\quad q = [1-\beta];\quad n_0 = [1-n-\alpha-\beta],\ \text{if } n+\alpha+\beta \leqq 1.
$$

Let N_j $(j = 1, 2, 3)$ denote the number of zeros of $J_n(x;\alpha,\beta)$ (assumed $\not\equiv 0$) in the intervals $(-\infty < x < -1)$, $(-1 < x < 1)$, $(1 < x < \infty)$ respectively, and K_i denote the number of zeros of $J_n(x;\alpha+i,\beta)$ ($\not\equiv 0$) in $(-\infty < x < -1)$, where the smallest and largest such zeros (if they exist) are λ_i, η_i respectively. Then, $K_0 = N_1$, $\lambda_i \leqq \eta_i < -1$. By (2, 10),

$$
(11)\qquad (n-i)(n+i+\alpha+\beta-1)\gamma_i + 2(i+1)(i+\alpha)\gamma_{i+1} = 0 \qquad (i = 0, 1, \cdots, n-1).
$$

Substituting $J_n(x;\alpha,\beta) = (1+x)^{1-\alpha}g(x)$ in (2), we obtain

$$
(12)\qquad
\begin{aligned}
J_n(x;1-p,\beta) &= (-1)^p\cdot p!\binom{n+\beta-1}{p}(x+1)^pJ_{n-p}(x;1+p,\beta) \qquad (n \geqq p),\\
J_n(x;\alpha,1-q) &= (-1)^q\cdot q!\binom{n+\alpha-1}{q}(x-1)^qJ_{n-q}(x;\alpha,1+q) \qquad (n \geqq q),\\
J_n(x;1-p,1-q) &= (-1)^{p+q}(p+q)!\binom{n}{p+q}(x+1)^p(x-1)^q\\
&\quad\cdot J_{n-p-q}(x;1+p,1+q) \qquad (n \geqq p+q).
\end{aligned}
$$

These formulas show the effect produced by the indicated changes in α, β. It is interesting to observe that the orthogonality relation (3) (derived for $\alpha, \beta > 0$) is still true if α, β are non-positive integers provided $m, n \geqq p+q$; for example, using (12), we have, by (3),

$$
\begin{aligned}
&\int_{-1}^{1}(1+x)^{-p}(1-x)^{-q}J_m(x;1-p,1-q)J_n(x;1-p,1-q)\,dx\\
&= [(p+q)!]^2\binom{m}{p+q}\binom{n}{p+q}\int_{-1}^{1}(1+x)^p(1-x)^qJ_{m-p-q}(x;1+p,1+q)\\
&\quad\cdot J_{n-p-q}(x;1+p,1+q)\,dx = 0 \qquad (m, n \geqq p+q).
\end{aligned}
$$

Formulas (2, 6, 10, 12) yield the following properties of $J_n(x; \alpha, \beta)$: (i) it $\equiv 0$ if, and only if, $n + n_0 + \alpha + \beta = 1$, $\alpha + p = 1$, $n \geqq p > n_0$; (ii) it has no multiple zeros, except possibly at $x = \pm 1$; (iii) it has a zero of multiplicity p at $x = -1$ if, and only if, $\alpha + p = 1$, $n \geqq p$; (iv) it has a zero of multiplicity q at $x = 1$ if, and only if, $\beta + q = 1$, $n \geqq q$; (v) it has a zero of multiplicity $[n - n_0]$ at infinity if, and only if, $n + n_0 + \alpha + \beta = 1$.

In what follows we may assume $n > 0$ and $J_n(x; \alpha, \beta) \not\equiv 0$.

2. For the further study of zeros of $J_n(x)$ we make use of the following lemma [10]:

Lemma 1. *If $f(x) \not\equiv 0$ satisfies the differential equation $r(x)f''(x) + s(x)f'(x) + t(x)f(x) = 0$ in a certain interval (c, d), where $r(x)$, $s(x)$, $t(x)$ are continuous and $r(x) \cdot t(x) < 0$, then $f(x)$ can have at most one zero inside (c, d).*

We now consider two cases.

Case 1. $n + \alpha + \beta > 1$.

Theorem 1. *If $n + \alpha + \beta > 1$, then $N_2 = [n - p - q]$ and N_1, N_3 are each either 0 or 1. Furthermore, if $\alpha + p = 1$, $n \geqq p$, then $N_1 = 0$; otherwise, $N_1 \equiv \min(n, p) \pmod 2$.*[2] *If $\beta + q = 1$, $n \geqq q$, then $N_3 = 0$; otherwise, $N_3 \equiv \min(n, q) \pmod 2$.*

Proof. N_j $(j = 1, 3)$ is 0 or 1 by Lemma 1.

(i) $n \leqq p$. (a) $n < p$ or $n = p$ and $\alpha \not\equiv r$. By (10, 11), $\gamma_i\gamma_{i+1} > 0$ $(i = 0, 1, \cdots, n - 1)$. It follows from Descartes' rule of signs that $N_2 = N_3 = 0$ and $N_1 \equiv n \pmod 2$.

(b) $n = p$, $\alpha + p = 1$. From (12), $N_1 = N_2 = N_3 = 0$.

(ii) $n \leqq q$. This reduces to (i) by means of (6).

(iii) $n > p$, $n > q$. Hence, $n \geqq p + q$. Lawton's method[3] shows $N_2 = n - p - q$. (6, 10, 11) enable us to complete the proof as in (i).

Case 2. $n + \alpha + \beta \leqq 1$. It suffices to illustrate the method for $0 < n + n_0 + \alpha + \beta < 1$, $\alpha \not\equiv r$. By Theorem 1,

$$K_{n_0+1} \text{ is } 0 \text{ or } 1 \qquad \text{and} \equiv \min(n, [p - n_0 - 1]) \pmod 2. \tag{13}$$

Let i be an integer such that $1 \leqq i \leqq n_0 + 1$. If $K_i \geqq 1$, it follows from (8, 9): (i) $J_n(x; \alpha + i - 1, \beta)$ and $J_n(x; \alpha + i, \beta)$ have no common zero inside $(-\infty, -1)$, (ii) $J_n(x; \alpha + i - 1, \beta)$ has exactly $K_i - 1$ zeros which separate the K_i zeros of $J_n(x; \alpha + i, \beta)$ inside $(-\infty, -1)$, (iii) $J_n(x; \alpha + i - 1, \beta)$ has at most one zero less than λ_i and at most one zero inside $(\eta_i, -1)$. Furthermore, making use of (1) and of the fact that $J_n(x)$ is a polynomial, we get:

[2] $c \equiv \min(a, b) \pmod 2$ means $c \equiv d \pmod 2$, where $d = \min(a, b)$.

[3] For $n \geqq p + q + 1$, Lawton proved that $J_n(x; \alpha, \beta)$ has (i) exactly $n - p - q$ zeros inside $(-1, 1)$, (ii) a zero of multiplicity p at $x = -1$ if $\alpha + p = 1$, (iii) a zero of multiplicity q at $x = 1$ if $\beta + q = 1$.

$$K_{i-1} = K_i + 1 \ (i > p), \quad K_i \ (i \leqq p < n + i), \quad |K_i - 1| \ (n + i \leqq p) \qquad (1 \leqq i \leqq n_0 + 1 \leqq n),$$

$$\text{(14)} \quad K_{i-1} = K_i \ (i > p), \quad |K_i - 1| \ (i \leqq p < n + i), \quad K_i \ (n + i \leqq p) \qquad (1 \leqq i \leqq n_0 - n + 1, n \leqq n_0),$$

$$K_{i-1} = K_i + 1 \ (i > p), \quad K_i \ (i \leqq p + n + i), \quad |K_i - 1| \ (n + i \leqq p) \qquad (n_0 - n + 2 \leqq i \leqq n_0 + 1, n \leqq n_0).$$

(14) remains valid if $K_i = 0$. By a repeated application of (14) and (13), we derive the following results:

(i) $n \geqq n_0 + 1$. (a) $p \geqq n_0 + 1$. N_1 is 0 or 1 and $\equiv n_0 + 1 - \min (n, p) \pmod 2$. (b) $p < n_0 + 1$. $N_1 = n_0 - p + 1$.

(ii) $n \leqq n_0$. $N_1 = [n - p]$.

We summarize our conclusions in

THEOREM 2. (i) $0 < n + n_0 + \alpha + \beta < 1$. *If* $n \geqq n_0 + 1$, $n_0 + 1 < p$ *and* $\alpha \neq r$, *or* $\alpha + p = 1$, $n < p$, *then* N_1 *is* 0 *or* 1 *and* $\equiv n_0 + 1 - \min (n, p) \pmod 2$. *If* $n \geqq p > n_0 + 1$, $\alpha + p = 1$, *then* $N_1 = 0$. *If* $n \geqq n_0 + 1 \geqq p$, *then* $N_1 = n_0 - p + 1$. *If* $n \leqq n_0$, *then* $N_1 = [n - p]$. (ii) $n + n_0 + \alpha + \beta = 1$. *Here,* $N_1 = [\min (n, n_0) - p]$.

THEOREM 3. *If* $n + \alpha + \beta \leqq 1$, *then* N_2 *is* 0 *or* 1.

Proof. If $n + \alpha + \beta < 1$, this follows from Lemma 1; if $n + \alpha + \beta = 1$, we use (2).

Analogous results regarding N_3 are obtained (see (6)) by interchanging α and β, p and q. Theorem 2 and the corresponding one for N_3 yield

COROLLARY. (i) *If* $0 < n + n_0 + \alpha + \beta < 1$, *it is impossible to have both* $N_1 > 1$, $N_3 > 1$; *if, in addition,* $\alpha = \beta$, *then* $N_1 = N_3 = 1$. (ii) *If* $n + n_0 + \alpha + \beta = 1$, *then at least one of the numbers* N_1, N_3 *is zero*; *if, in addition,* $\alpha = \beta$, *then* $N_1 = N_3 = 0$.

If $n + \alpha + \beta \leqq 1$, we have determined the number of real zeros of $J_n(x; \alpha, \beta)$, which do not lie inside $(-1, 1)$. Hence, we may find N_2 from Theorem 3, since the total number of real zeros is even or odd according as n is even or odd. Illustrations, such as the one below, show that the conclusions of the above theorems and corollary can not be improved. We note that these results differ from the case of Laguerre polynomials [10] in that we may have either N_1 or N_3 greater than unity.

Illustration: $n = 3$, $\alpha = -\dfrac{3}{4}$, $\beta = -\dfrac{7}{2}$. Here, $p = 1$, $q = 4$, $n_0 = 2$,

$$J_3(x) = \frac{1}{2^6} \{3^2 \cdot 5(x + 1)^3 + 2 \cdot 3^3 \cdot 5^2(x + 1)^2 + 2^2 \cdot 3^3 \cdot 5(x + 1) - 2^3 \cdot 3 \cdot 5\}.$$

Since the discriminant of the equation $J_3(x) = 0$ is positive, all of its roots are real; thus, $N_1 = 2$, $N_2 = 1$, $N_3 = 0$, in accordance with the preceding theorems.

3. Hereafter, our analysis is confined to the study of the zeros $x_{i,n}$ (all inside $(-1, 1)$) of $J_n(x; \alpha, \beta)$ for the case $\alpha, \beta > 0$. It is known [11] that

$$x_{i,n}(\alpha, \beta) < x_{i,n}(\alpha + 1, \beta) < x_{i+1,n}(\alpha, \beta) \qquad (i = 1, 2, \cdots, n - 1);$$

furthermore, by (8, 9),

$$\begin{aligned}(n + \alpha + \beta)(x + 1)J_n(x; \alpha + 2, \beta) &= [(2n + \alpha + \beta)(x + 1) + 2\alpha] \\ &\cdot J_n(x; \alpha + 1, \beta) - 2(n + \alpha)J_n(x; \alpha, \beta).\end{aligned} \tag{15}$$

Thus, we obtain the following interesting inequalities, to be used later:

$$\begin{gathered} x_{i,n}(\alpha, \beta) < x_{i,n}(\alpha + 1, \beta) < x_{i,n}(\alpha + 2, \beta) < x_{i+1,n}(\alpha, \beta), \\ x_{i,n}(\alpha, \beta) < x_{i+1,n}(\alpha, \beta + 2) < x_{i+1,n}(\alpha, \beta + 1) < x_{i+1,n}(\alpha, \beta), \\ (i = 1, 2, \cdots, n - 1). \end{gathered} \tag{16}$$

By means of (10), we see that $x_{i,n}(\alpha, \beta) \to -1 + \dfrac{2k}{k + 1} = \dfrac{k - 1}{k + 1}$ if $\beta \to \infty$, $\dfrac{\alpha}{\beta} \to k$ ($\geqq 0$, finite or infinite). In particular, for $k = \infty$,

$$x_{i,n}(\alpha, \beta) \to 1 \text{ if } \alpha \to \infty \qquad (i = 1, 2, \cdots, n).$$

An interesting relation for the $x_{i,n}$ may be obtained by putting $\bar{\phi}_n(x; \alpha, \beta) \equiv \bar{\phi}_n(x) \equiv x^n + \cdots$, where

$$\int_0^1 x^{\alpha-1}(1 - x)^{\beta-1}\bar{\phi}_m(x)\bar{\phi}_n(x)dx = 0 \qquad (m \neq n; m, n = 0, 1, \cdots).$$

If we make the substitution $x = (1 + y)/2$, thus reducing the interval $(0, 1)$ to $(-1, 1)$, and compare with the orthogonality relation (3) for $\phi_n(x; \alpha, \beta)$, we find that $\phi_n(x)$ and $\bar{\phi}_n[(1 + x)/2]$ differ by a constant factor only. Thus

$$\phi_n(x) \equiv 2^n\bar{\phi}_n\left(\frac{1 + x}{2}\right), \qquad \bar{x}_{i,n} \equiv \frac{1 + x_{i,n}}{2},$$

where $\bar{x}_{i,n} \equiv \bar{x}_{i,n}(\alpha, \beta)$ are the zeros of $\bar{\phi}_n(x; \alpha, \beta)$. Making use of the known relations [11]:

$$\begin{aligned} x_{i,2n}^2(\beta, \beta) &= \bar{x}_{i-n,n}(\tfrac{1}{2}, \beta) \qquad && (n + 1 \leqq i \leqq 2n), \\ x_{i,2n+1}^2(\beta, \beta) &= \bar{x}_{i-n-1,n}(\tfrac{3}{2}, \beta) && (n + 2 \leqq i \leqq 2n + 1), \end{aligned}$$

we get

$$\begin{gathered} 1 + x_{i,n}(\tfrac{1}{2}, \beta) = 2x_{n+i,2n}^2(\beta, \beta), \qquad 1 + x_{i,n}(\tfrac{3}{2}, \beta) = 2x_{n+i+1,2n+1}^2(\beta, \beta) \\ (i = 1, 2, \cdots, n), \end{gathered} \tag{17}$$

interesting relations connecting the symmetric case ($\alpha = \beta$) with two particular non-symmetric cases corresponding to the same interval $(-1, 1)$. We shall make important use of (17).

With W. Hahn, let

$$p_1 = \sum_{i=1}^{n} \frac{1}{1 + x_{i,n}}, \qquad p_2 = \sum_{i=1}^{n} \frac{1}{(1 + x_{i,n})^2}, \qquad \frac{1}{p_1} < 1 + x_{1,n} < \frac{p_1}{p_2}.$$

Evaluating p_1, p_2 by means of (10) and using (6), we get

$$\frac{2\alpha}{n(n+\alpha+\beta-1)} < 1 + x_{1,n} < \frac{2\alpha(\alpha+1)}{n(n+\alpha+\beta-1)+\alpha(\alpha+\beta)},$$

$$\text{(18)} \quad \frac{2\beta}{n(n+\alpha+\beta-1)} < 1 - x_{n,n} < \frac{2\beta(\beta+1)}{n(n+\alpha+\beta-1)+\beta(\alpha+\beta)},$$

$$x_{1,n} = -1 + O\left(\frac{1}{n^2}\right), \qquad x_{n,n} = 1 + O\left(\frac{1}{n^2}\right) \qquad (\alpha, \beta > 0, n \to \infty).$$

These results may be greatly improved for some special α, β (see (32)).[4] We note also that

$$1 + x_{1,n} = \alpha\left[\frac{2}{n(n+\beta-1)} + o(1)\right] \qquad (\alpha \to 0),$$

$$1 - x_{n,n} = \beta\left[\frac{2}{n(n+\alpha-1)} + o(1)\right] \qquad (\beta \to 0),$$

$$x_{1,n}(0, \beta) = -1, \qquad x_{n,n}(\alpha, 0) = 1$$

(see (12)).

From (17) and (18) we obtain

$$\text{(19)} \quad \begin{aligned} &\frac{\sqrt{2}}{2\sqrt{m(m+\alpha-\frac{1}{2})}} < -x_{m,2m} = x_{m+1,2m} < \frac{\sqrt{3}}{2\sqrt{m(m+\alpha-\frac{1}{2})+\frac{1}{2}(\alpha+\frac{1}{2})}} \\ &\frac{\sqrt{6}}{2\sqrt{m(m+\alpha+\frac{1}{2})}} < -x_{m,2m+1} = x_{m+2,2m+1} < \frac{\sqrt{15}}{2\sqrt{m(m+\alpha+\frac{1}{2})+\frac{3}{2}(\alpha+\frac{3}{2})}} \end{aligned}$$

$$(\alpha = \beta > 0).$$

In some cases these results may be improved (see (29)).

4. We proceed along the lines of E. R. Neumann and Winston to obtain bounds for the general $x_{i,n}$. (5, 6, 8, 9) yield

$$\text{(20)} \quad \frac{1}{n+1} J'_{n+1}(x) = -\frac{(2n+\alpha+\beta)(x+x_0)}{n+\alpha+\beta-1} J'_n(x) - (2n+\alpha+\beta)J_n(x) \qquad \left(x_0 \equiv \frac{\alpha-\beta}{2n+\alpha+\beta}\right),$$

$$\text{(21)} \quad \frac{(2n+\alpha+\beta)(1-x^2)}{n+1} J'_{n+1}(x) = -(2n+\alpha+\beta)(x+x_0)J_{n+1}(x) - 4(n+\alpha)(n+\beta)J_n(x),$$

[4] It is to be noted that these results were obtained before the publication of Buell's [14] paper which supplements but does not supplant formulas (18), (19), etc. Formulas (31), (32), etc. were not given by Buell.

$$J_{n+1}(x; \alpha - 1, \beta - 1) = (1 - x^2)J'_n(x; \alpha, \beta) + [\alpha - \beta + (\alpha + \beta - 2)x]J_n(x; \alpha, \beta). \tag{22}$$

Let $\mu_{i,n}(\alpha, \beta) \equiv \mu_{i,n}$ $(i = 1, 2, \cdots, n - 1)$ be the zeros of $J'_n(x; \alpha, \beta)$. In view of (5), it is known that

$$\mu_{i,n}(\alpha, \beta) = x_{i,n-1}(\alpha + 1, \beta + 1), \qquad \begin{Bmatrix} x_{i,n+1} \\ \mu_{i-1,n} \end{Bmatrix} < \begin{Bmatrix} x_{i,n} \\ \mu_{i,n+1} \end{Bmatrix} < \begin{Bmatrix} x_{i,n+1} \\ \mu_{i,n} \end{Bmatrix},$$

where either of the terms in each bracket may be used. It may be shown by induction on i that $J_n(x_{i,n+1}) \cdot J_{n+1}(x_{i,n}) < 0$, so that (20), (21) give

$$x_{i,n} \gtreqless \mu_{i,n+1} \text{ according as } x_{i,n} \gtreqless -x_0,$$

$$x_{i+1,n+1} \gtreqless \mu_{i,n} \text{ according as } x_{i+1,n+1} \lesseqgtr x_0.$$

We see from (2) that there is no point of inflection of the curve $y = J_n(x)$ inside $(x_{i,n}, \mu_{i,n})$ if $x_{i,n} \leqq \dfrac{\alpha - \beta}{\alpha + \beta}$, and no such point inside $(\mu_{i,n}, x_{i+1,n})$ if $x_{i,n} \geqq \dfrac{\alpha - \beta}{\alpha + \beta}$. By (20), (21), (22), let

$$I_1 \equiv (2n + \alpha + \beta) \int_{x_{i,n}}^{x_{i+1,n}} J_n(x)dx$$

$$= \frac{n + \alpha + \beta - 1}{(n + 1)(n + \alpha + \beta - 2)} [J_{n+1}(x_{i,n}) - J_{n+1}(x_{i+1,n})],$$

$$|I_1| > \frac{n + \alpha + \beta - 1}{(n + 1)(n + \alpha + \beta - 2)} |J_{n+1}(x_{i,n})|.$$

If $x_{i,n} \leqq \dfrac{\alpha - \beta}{\alpha + \beta}$, then $\dfrac{|I_1|}{2n + \alpha + \beta}$ is less than the area of the triangle formed by the x-axis, the tangent to $y = J_n(x)$ at $x = x_{i,n}$, and the perpendicular to the x-axis at $x = x_{i+1,n}$, so that by (20, 21),

$$|I_1| < \frac{(2n + \alpha + \beta)(x_{i+1,n} - x_{i,n})^2}{2} |J'_n(x_{i,n})|$$

$$= \frac{(n + \alpha + \beta - 1)(x_{i+1,n} - x_{i,n})^2}{2(1 - x_{i,n}^2)} |J_{n+1}(x_{i,n})|.$$

Hence $\left(\text{similarly if } x_{i,n} \geqq \dfrac{\alpha - \beta}{\alpha + \beta}\right)$,

$$\begin{aligned} x_{i+1,n} - x_{i,n} &> \left[\frac{2(1 - x_{i,n}^2)}{(n + 1)(n + \alpha + \beta - 2)}\right]^{\frac{1}{2}}, & x_{i,n} &\leqq \frac{\alpha - \beta}{\alpha + \beta}, \\ x_{i+1,n} - x_{i,n} &> \left[\frac{2(1 - x_{i+1,n}^2)}{(n + 1)(n + \alpha + \beta - 2)}\right]^{\frac{1}{2}}, & x_{i,n} &\geqq \frac{\alpha - \beta}{\alpha + \beta}. \end{aligned} \tag{23}$$

The inequalities (23) lead to the following bounds for $x_{i,n}$:

$$(24)\quad \begin{aligned} 1 + x_{i,n} &> \frac{(i + 2\sqrt{\alpha_2} - 1)^2}{4(n+1)(n+\alpha+\beta-2)}\cdot\frac{\alpha\beta_1}{\alpha_2}, \qquad x_{i,n} \leqq \frac{\alpha-\beta}{\alpha+\beta}, \quad n \geqq 3,\\ 1 + x_{i,n} &< 2 - \frac{(n - i + 2\sqrt{\beta_2})^2}{4(n+1)(n+\alpha+\beta-2)}\cdot\frac{\alpha_1\beta}{\beta_2}, \quad x_{i,n} \geqq \frac{\alpha-\beta}{\alpha+\beta}, \quad n \geqq 3. \end{aligned}$$

$$\alpha_1 = \min\left(1, \frac{2\alpha}{\alpha+\beta}\right), \qquad \beta_1 = \min\left(1, \frac{2\beta}{\alpha+\beta}\right), \qquad \alpha_2 = \max(\alpha, \omega),$$

$$\beta_2 = \max(\beta, \omega), \qquad \omega \equiv \left(\frac{1+\sqrt{2}}{4}\right)^2 \equiv 0.364 \cdots.$$

It suffices to outline the proof for the first inequality (24), when $\alpha \geqq \beta$.

Case 1. $\alpha \geqq \omega$. Since $x_{i,n} \leqq \dfrac{\alpha-\beta}{\alpha+\beta}$, it follows that $1 - x_{i,n} \geqq 2\beta/(\alpha+\beta)$ with $0 < 2\beta/(\alpha+\beta) \leqq 1$. Hence, by (23),

$$(1 + x_{i+1,n}) - (1 + x_{i,n}) > \left[\frac{2(1+x_{i,n})}{(n+1)(n+\alpha+\beta-2)}\cdot\frac{2\beta}{\alpha+\beta}\right]^{\frac{1}{2}},$$

$$\rho_{i+1} > \rho_i + \sqrt{2\rho_i}, \qquad \rho_i \equiv \frac{\alpha+\beta}{2\beta}(n+1)(n+\alpha+\beta-2)(1+x_{i,n}).$$

Following E. R. Neumann, we prove (by induction) that

$$\rho_i > \sigma_i \equiv \frac{(i + 2\sqrt{\alpha} - 1)^2}{4},$$

at least if $n \geqq 3$. Thus,

$$1 + x_{i,n} > \frac{(i + 2\sqrt{\alpha} - 1)^2}{4(n+1)(n+\alpha+\beta-2)}\cdot\frac{2\beta}{\alpha+\beta} \qquad (n \geqq 3).$$

Case 2. $0 < \alpha < \omega$. Here, we prove that

$$\rho_i \equiv \frac{\omega}{\alpha}\cdot\frac{\alpha+\beta}{2\beta}(n+1)(n+\alpha+\beta-2)(1+x_{i,n}) > \sigma_i \equiv \frac{(i + 2\sqrt{\omega} - 1)^2}{4},$$

$$1 + x_{i,n} > \frac{(i + 2\sqrt{\omega} - 1)^2}{4(n+1)(n+\alpha+\beta-2)}\cdot\frac{\alpha}{\omega}\cdot\frac{2\beta}{\alpha+\beta} \qquad (n \geqq 3).$$

Continuing with E. R. Neumann's method, we may, by means of a complicated analysis, find certain upper and lower bounds for the negative $x_{i,n}$ in the symmetric case ($\alpha = \beta$) as well as in the non-symmetric case. Instead of giving these results, we proceed to develop a method which yields much more.

5. Since $x_{i,n}(\alpha, \beta)$ for $\alpha, \beta = 1/2, 3/2$ are known [11] explicitly, Markoff's theorem[5] immediately gives

[5] $\dfrac{\partial x_{i,n}(\alpha,\beta)}{\partial\alpha} > 0$, $\dfrac{\partial x_{i,n}(\alpha,\beta)}{\partial\beta} < 0$ $(\alpha \neq \beta)$; $\dfrac{\partial x_{i,n}(\alpha,\alpha)}{\partial\alpha} \gtrless 0$ according as $x_{i,n} \lessgtr 0$.

$$(25)\quad \begin{gathered} -\cos\frac{2i-1}{2n+1}\pi \leqq x_{i,n}(\alpha,\beta) \leqq -\cos\frac{2i}{2n+1}\pi \qquad (\tfrac{1}{2} \leqq \alpha, \beta \leqq \tfrac{3}{2}; i = 1, 2, \cdots, n), \\ x_{i,n}(\alpha,\beta) = -\cos\frac{2i}{2n+1}\pi + O\left(\frac{i}{n^2}\right). \end{gathered}$$

We may generalize this result for all α, β by means of (16), Markoff's theorem, and [11]

$$(26)\quad x_{1,n+1}(\alpha,\beta) < x_{1,n}(\alpha,\beta+1), \qquad x_{n,n}(\alpha+1,\beta) < x_{n+1,n+1}(\alpha,\beta).$$

For example, if $\alpha, \beta \geqq 1/2$, there exist non-negative integers r, s such that

$$\tfrac{1}{2} \leqq \alpha - 2r - \sigma_1 < \tfrac{3}{2}, \qquad \tfrac{1}{2} \leqq \beta - 2s - \sigma_2 < \tfrac{3}{2} \qquad (\sigma_1, \sigma_2 = 0, 1).$$

If $s + 1 < i \leqq n - r - 1$, then

$$\begin{aligned} &x_{i-s-1,n}(\alpha - 2r - \sigma_1, \beta - 2s - \sigma_2) \leqq x_{i-s-1,n}(\alpha, \beta - 2s - \sigma_2) \\ &< x_{i-s,n}(\alpha, \beta - 2s) \leqq x_{i,n}(\alpha,\beta) \leqq x_{i+r,n}(\alpha - 2r, \beta) < x_{i+r+1,n}(\alpha - 2r - \sigma_1, \beta) \\ &\qquad \leqq x_{i+r+1,n}(\alpha - 2r - \sigma_1, \beta - 2s - \sigma_2). \end{aligned}$$

Hence, for all $\alpha, \beta \geqq 1/2$,

$$(27)\quad \begin{gathered} -\cos\frac{2i-2s-3}{2n+1}\pi \leqq x_{i,n}(\alpha,\beta) \leqq -\cos\frac{2i+2r+2}{2n+1}\pi \\ (\tfrac{1}{2} \leqq \alpha - 2r - \sigma_1 < \tfrac{3}{2}, \tfrac{1}{2} \leqq \beta - 2s - \sigma_2 < \tfrac{3}{2}; \sigma_1, \sigma_2 = 0, 1; \\ s + 1 < i \leqq n - r - 1). \end{gathered}$$

It follows that

$$(28)\quad x_{i,n}(\alpha,\beta) \to -\cos t\pi \qquad \text{if } \frac{i}{n} \to t \text{ as } n \to \infty \qquad (\alpha, \beta > 0).$$

Markoff's theorem and (16) give, further,

$$(29)\quad \begin{aligned} &\frac{\sqrt{2}}{2(m + \alpha - \frac{1}{2})} < x_{m+1,2m}(\alpha,\alpha) \leqq x_{m+1,2m}(\tfrac{1}{2},\tfrac{1}{2}) = \sin\frac{\pi}{4m} < \frac{\pi}{4m} \qquad (\alpha \geqq \tfrac{1}{2}), \\ &\sin\frac{\pi}{4m} = x_{m+1,2m}(\tfrac{1}{2},\tfrac{1}{2}) < x_{m+1,2m}(\alpha,\alpha) < x_{m+1,2m}(\tfrac{1}{2},\alpha) \\ &\qquad < x_{m+2,2m}(\tfrac{1}{2},\tfrac{1}{2}) = \sin\frac{3\pi}{4m} < \frac{3\pi}{4m} \qquad (0 < \alpha < \tfrac{1}{2}), \\ &\frac{\sqrt{6}}{2(m + \alpha + \frac{1}{2})} < x_{m+2,2m+1}(\alpha,\alpha) \leqq x_{m+2,2m+1}(\tfrac{1}{2},\tfrac{1}{2}) = \sin\frac{\pi}{2m+1} < \frac{\pi}{2m+1} \\ &\qquad (\alpha \geqq \tfrac{1}{2}), \\ &\sin\frac{\pi}{2m+1} = x_{m+2,2m+1}(\tfrac{1}{2},\tfrac{1}{2}) < x_{m+2,2m+1}(\alpha,\alpha) < x_{m+2,2m+1}(\tfrac{1}{2},\alpha) \\ &\qquad < x_{m+3,2m+1}(\tfrac{1}{2},\tfrac{1}{2}) = \sin\frac{2\pi}{2m+1} < \frac{2\pi}{2m+1} \qquad (0 < \alpha < \tfrac{1}{2}). \end{aligned}$$

If $\alpha \geqq \frac{1}{2}$, the upper bounds in (29) are better than in (19). Moreover,

$$
\begin{aligned}
\sin\left(\frac{2i-1+\sigma}{2m+1+\sigma}\cdot\frac{\pi}{2}\right) &= x_{m+i+\sigma,2m+\sigma}(\tfrac{3}{2},\tfrac{3}{2}) \leqq x_{m+i+\sigma,2m+\sigma}(\alpha,\alpha) \\
&\leqq x_{m+i+\sigma,2m+\sigma}(\tfrac{1}{2},\tfrac{1}{2}) = \sin\frac{2i-1+\sigma}{2m+\sigma}\cdot\frac{\pi}{2} \qquad (30)\\
&\qquad (\tfrac{1}{2} \leqq \alpha \leqq \tfrac{3}{2};\ \sigma = 0, 1;\ i = 1, 2, \cdots, m).
\end{aligned}
$$

We thus obtain the following asymptotic relations:

$$
\begin{aligned}
x_{m+i+\sigma,2m+\sigma}(\alpha,\alpha) &= \sin\left(\frac{2i-1+\sigma}{2m+1+\sigma}\cdot\frac{\pi}{2}\right) + O\left(\frac{i}{m^2}\right) \\
&\qquad (\tfrac{1}{2} \leqq \alpha \leqq \tfrac{3}{2};\ \sigma = 0, 1;\ i > 0), \qquad (31)\\
\frac{2m+1+\sigma}{2i-1+\sigma}\, x_{m+i+\sigma,2m+\sigma}(\alpha,\alpha) &\to \frac{\pi}{2} \quad \text{if } \frac{i}{m} \to 0 \quad \text{as } m \to \infty.
\end{aligned}
$$

Making use of (17), we have

$$
\begin{aligned}
\left[\frac{2n+1+\sigma}{2i-1+\sigma}\right]^2 [1 + x_{i,n}(\tfrac{1}{2}+\sigma,\beta)] &\to \frac{\pi^2}{2} \quad \text{if } \frac{i}{n} \to 0 \quad \text{as } n \to \infty \\
&\qquad (\tfrac{1}{2} \leqq \beta \leqq \tfrac{3}{2};\ \sigma = 0, 1), \qquad (32)\\
x_{1,n}(\tfrac{1}{2},\beta) = -1 + \frac{\pi^2 + o(1)}{8n^2}, \quad x_{1,n}(\tfrac{3}{2},\beta) &= -1 + \frac{\pi^2 + o(1)}{2n^2} \quad (\tfrac{1}{2} \leqq \beta \leqq \tfrac{3}{2}),\\
x_{n,n}(\alpha,\tfrac{1}{2}) = 1 + \frac{\pi^2 + o(1)}{8n^2}, \quad x_{n,n}(\alpha,\tfrac{3}{2}) &= 1 + \frac{\pi^2 + o(1)}{2n^2} \quad (\tfrac{1}{2} \leqq \alpha \leqq \tfrac{3}{2}).
\end{aligned}
$$

In view of (16, 17, 29, 32), we have

$$
(33) \qquad \left[\frac{n}{i}\right]^2 [1 + x_{i,n}(\alpha,\beta)] \to \frac{\pi^2}{2} \quad \text{if } i, n \to \infty, \quad \frac{i}{n} \to 0 \qquad (0 < \alpha, \beta \leqq \tfrac{3}{2}).
$$

Although derived for $\alpha, \beta \leqq \frac{3}{2}$, (33) *remains valid*, in view of (16), for all $\alpha, \beta > 0$ (as in the case of (28)).

6. We apply the foregoing bounds for the zeros $x_{i,n}$ to the mechanical quadratures coefficients $H_{i,n}(\alpha, \beta) \equiv H_{i,n}$ (see (4)) by means of the Tchebycheff inequality

$$
(34) \quad 0 < H_{i,n} < \int_{x_{i-1,n}}^{x_{i+1,n}} p(x)\,dx \qquad (x_{0,n} \equiv -1,\ x_{n+1,n} \equiv 1;\ i = 1, 2, \cdots, n).
$$

It is known [9] that $H_{i,n}(\alpha, \beta) = H_{n-i+1,n}(\beta, \alpha)$. Also, by a method similar to that which leads to (17), we get

$$
(35) \quad
\begin{aligned}
H_{i,n}(\tfrac{1}{2},\beta) &= 2^{\beta+\frac{1}{2}} H_{n+i,2n}(\beta,\beta),\\
H_{i,n}(\tfrac{3}{2},\beta) &= 2^{\beta+\frac{3}{2}} x^2_{n+i+1,2n+1}(\beta,\beta) H_{n+i+1,2n+1}(\beta,\beta) \qquad (i = 1, 2, \cdots, n).
\end{aligned}
$$

In what follows, we obtain an upper bound of the order $\frac{1}{n^{2\alpha}}$ for all $H_{i,n}$ such that $1 < i \leqq C$ (arbitrarily fixed constant) in which case Winston [9] gave a lower bound of the order $\frac{1}{n^{2\alpha}}$ if $\alpha > \frac{1}{2}$. *Hence, the true order, with respect to n, of $H_{i,n}$ is $\frac{1}{n^{2\alpha}}$ $(\alpha > \frac{1}{2}, 1 < i \leqq C)$.* It is sufficient to illustrate the procedure in the special case $\frac{1}{2} \leqq \alpha, \beta < 1$. We have (by (34)),

$$H_{i,n} < (x_{i+1,n} - x_{i-1,n})(1 + x_{i-1,n})^{\alpha-1}(1 - x_{i+1,n})^{\beta-1}.$$

(Using the fact that $p(x)$ is decreasing for $-1 < x < \frac{\alpha - \beta}{\alpha + \beta - 2}$ and increasing for $\frac{\alpha - \beta}{\alpha + \beta - 2} < x < 1$ does not give any essential improvement.) By (25),

$$H_{i,n} < 2^{\alpha+\beta-1} \sin\left[\frac{4i-1}{2n+1}\cdot\frac{\pi}{2}\right] \sin\left[\frac{5}{2n+1}\cdot\frac{\pi}{2}\right]\left[\sin\frac{2i-3}{2n+1}\cdot\frac{\pi}{2}\right]^{2(\alpha-1)}$$

$$\cdot\left[\cos\frac{2i+2}{2n+1}\cdot\frac{\pi}{2}\right]^{2(\beta-1)} = O\left(\frac{1}{n^{2\alpha-\epsilon(2\alpha-1)}}\right), \quad \text{if } i = O(n^{\epsilon}) \qquad (0 \leqq \epsilon < 1).$$

Likewise, if $0 < \alpha = \beta < \frac{1}{2}$, then $H_{m+i,2m+\sigma} = O\left(\frac{1}{m}\right)$, if $1 \leqq i \leqq C$ $(\sigma = 0, 1)$. Since Winston gave a lower bound for $H_{i,n}$ of the same order, *this is its true order.*

$H_{1,n}$ and $H_{n,n}$ may be treated as follows. ((34) gives an upper bound in case $\alpha \geqq 1$ but $p(x) \to \infty$ as $x \to -1$ for $0 < \alpha < 1$.) For $\alpha > \frac{1}{2}$ and n sufficiently large, Winston showed that $H_{1,n} < H_{2,n}$. He gave also an upper bound of the order $\frac{1}{n^{2\alpha}}$ for $H_{1,n}$ if $\alpha \leqq \frac{1}{2}$, except when $\beta < \alpha$ in which case $H_{1,n} < H_{2,n}$, so that $H_{1,n} = O\left(\frac{1}{n^{2\alpha}}\right)$. On the other hand, if $\varphi_n(x)$ is the normalized Jacobi polynomial of degree n,

$$H_{1,n} = \frac{1}{K_n(x_{1,n})} > \frac{1}{K_n(-1)} = \frac{2^{\alpha+\beta-1}\,\Gamma(\alpha)\Gamma(\alpha+1)\Gamma(n+1)\Gamma(n+\beta)}{\Gamma(n+\alpha+1)\,\Gamma(n+\alpha+\beta)} = O\left(\frac{1}{n^{2\alpha}}\right),$$

[11], since $K'_n(x) \equiv \left(\sum_{i=0}^{n}\varphi_i^2(x)\right)' < 0$ for $x < x_{1,n}$. *Hence, the true order of $H_{1,n}$ is $\frac{1}{n^{2\alpha}}$ for all $\alpha > 0$. Likewise, the true order of $H_{n,n}$ is $\frac{1}{n^{2\beta}}$ for all $\beta > 0$.*

The above upper bounds enable us, in case the behavior of $f(x)$ near $x = \pm 1$ is properly specified, to estimate how much the mechanical quadratures formula (4) is affected if we omit even infinitely many of the terms $H_{i,n}f(x_{i,n})$ corresponding to the $x_{i,n}$ near ± 1.

Illustration. Consider $f(x)$ such that

$$|f(x)| = O\left(\frac{1}{(1-|x|)^l}\right) \qquad (0 < 1 - |x| < \delta;\ \delta,\ l \text{ positive constants}).$$

Then we can omit in (4) all terms $H_{i,n}f(x_{i,n})$ corresponding to

(36) $i < C_1 n^\epsilon$ or $i > n - C_2 n^\epsilon$ ($C_1 > 0$, $C_2 > 0$, ϵ given constants, $0 < \epsilon < 1$),

provided[6] $\beta \geqq \alpha$, $2\alpha - \epsilon(2\alpha - 1) > 2l$. In fact, for all such i, $|x_{i,n}| \to 1$ $(n \to \infty)$ (see (27)), so that, by (18), for n sufficiently large,

$$0 < 1 - |x_{i,n}| < \delta, \qquad 1 - |x_{i,n}| > \frac{2\alpha}{n(n + \alpha + \beta - 1)}.$$

Thus $|\sum H_{i,n}f(x_{i,n})|$, extended over all i as given in (36), is $= O\left(\frac{1}{n^{2\alpha-\epsilon(2\alpha-1)-2l}}\right)$ $\to 0$ as $n \to \infty$.

7. The method of the preceding section can be applied to Laguerre polynomials, in which case $p(x) = x^{\alpha-1}e^{-x}$ $(\alpha > 0)$ and the interval of orthogonality is $(0, \infty)$.

(i) $0 < \alpha \leqq 1$. Using (34) and the bounds for $x_{i,n}$ given by Winston, we get

$$H_{i,n} < e^{-\frac{(i+2\sqrt{\alpha_2}-2)^2}{4(n+1)}\cdot\frac{\alpha}{\alpha_2}} \cdot \left\{\frac{(i + 2\sqrt{\alpha_2} - 2)^2}{4(n+1)} \cdot \frac{\alpha}{\alpha_2}\right\}^\alpha$$
$$\cdot \left\{\frac{16(i + \alpha + 1)^2(n + 1)}{(i + 2\sqrt{\alpha_2} - 2)^2(n + \alpha)} \cdot \frac{\alpha_2}{\alpha} - 1\right\},$$
$$\alpha_2 = \max\left\{\alpha, \left(\frac{1 + \sqrt{2}}{4}\right)^2\right\} \qquad (i = 2, 3, \cdots, n - 1).$$

(ii) $\alpha > 1$.

$$H_{i,n} < e^{-\frac{(i+2\sqrt{\alpha}-2)^2}{4(n+1)}} \cdot \left\{\frac{4(i + \alpha + 1)^2}{n + \alpha}\right\}^\alpha \left\{1 - \frac{(i + 2\sqrt{\alpha} - 2)^2(n + \alpha)}{16(i + \alpha + 1)^2(n + 1)}\right\}$$
$$(i = 2, 3, \cdots, n - 1).$$

In particular,

$$H_{i,n} = O\left(\frac{1}{n^\alpha}\right) \qquad (1 < i \leqq C, \alpha > 0),$$

$$H_{i,n} = O(e^{-\gamma n}n^\alpha) \qquad \left(0 < C \leqq \frac{i}{n}, \gamma = \frac{C^2\alpha}{4\alpha_2}, \alpha > 0\right).$$

(Since $H_{n,n} < H_{n-1,n}$, at least if n is sufficiently large [9], this remains valid for

[6] This restriction is not essential, since we may always interchange α, β.

$i = n$.) Winston gave a lower bound of the order $\dfrac{1}{n^\alpha}$ for $H_{i,n}$ ($\alpha > \frac{1}{2}$; $i = 1, 2, \cdots, n$) and showed that[7]

$$H_{1,n} = O\left(\frac{1}{n^\alpha}\right) \quad (0 < \alpha \leqq \tfrac{1}{2}); \quad H_{1,n} < H_{2,n} \qquad (\alpha > \tfrac{1}{2}, n \text{ sufficiently large}).$$

Here, [11],

$$H_{1,n} > \frac{1}{K_n(0)} = \frac{\alpha\Gamma^2(\alpha)\Gamma(n+1)}{\Gamma(n+\alpha+1)} = O\left(\frac{1}{n^\alpha}\right) \qquad (\alpha > 0).$$

Hence, the true order of $H_{1,n}$ *is* $\dfrac{1}{n^\alpha}$ ($\alpha > 0$) *and the true order of* $H_{i,n}$ *is* $\dfrac{1}{n^\alpha}$ ($\alpha > \frac{1}{2}$, $1 \leqq i \leqq C$). We note, however, that $H_{i,n}$ is not of the same order with respect to n for all i.

In accordance with the results of Winston, similar results could be obtained for Hermite polynomials.

8. The above classical orthogonal polynomials of Jacobi (J), Laguerre (L), and Hermite (H) possess the remarkable property, important for the preceding discussion, that the derived polynomials are again orthogonal polynomials J, L, and H respectively, but with new parameters. For Hermite polynomials, we have a still more remarkable result: $\phi_n'(x) = n\phi_{n-1}(x)$, i.e., the weight function remains unchanged, or, in other words, *Hermite polynomials form a system of Appell* [13] *polynomials. We close by giving a simple proof that the Hermite polynomials are the only orthogonal polynomials with this property.*[8]

THEOREM 4. *If* $\{\phi_n(x) \equiv x^n + \cdots\}$, $n = 0, 1, \cdots$, *is an orthogonal system of polynomials, i.e.,* $\int_{-\infty}^{\infty} \phi_m(x)\phi_n(x)d\psi(x) = 0$ ($m \neq n$, $\psi(x)$ *monotone non-decreasing*), *such that* $\phi_n'(x) = n\phi_{n-1}(x)$ *then* $\{\phi_n(x)\}$ *is reducible, except for constant factors, to the Hermite system* $\{H_n(x)\}$ *of polynomials by means of a linear transformation on* x.

Proof.[9] Write with Appell [13]

$$(37) \qquad \phi_n(x) = x^n + \frac{n}{1}a_1x^{n-1} + \frac{n(n-1)}{2!}a_2x^{n-2} + \cdots + \frac{n}{1}a_{n-1}x + a_n,$$

where $a_1, a_2, \cdots, a_n, \cdots$ are given constants uniquely determined when

[7] Winston's result (corrected for a slight misprint) is

$$H_{i,n} \leqq \frac{\Gamma^2(\alpha)\Gamma(n)}{\Gamma(n+\alpha)} = O\left(\frac{1}{n^\alpha}\right) \qquad (0 < \alpha \leqq \tfrac{1}{2};\ i = 1, 2, \cdots, n).$$

[8] The same result has recently been derived by Meixner [12] as a consequence of more general considerations.

[9] The present proof was obtained before the proof recently published by Shohat [15].

$\{\phi_n(x)\}$ is given. It follows (by induction for $n > 1$) from the fundamental recurrence relation [11]

$$\phi_{n+1}(x) = (x - c_{n+1})\phi_n(x) - \lambda_{n+1}\phi_{n-1}(x) \qquad (n = 1, 2, \cdots)$$

that

$$\phi_n(-x - a_1) \equiv (-1)^n\phi_n(x - a_1) \qquad (n = 0, 1, \cdots).$$

Hence, $\phi_n(x - a_1)$ contains only even or only odd powers of x, so that

$$\begin{aligned} \phi_{2m+\sigma}(x - a_1) &= \sum_{i=0}^{m} C_{i,m}x^{2(m-i)+\sigma} \\ \phi_{2m+\sigma}(x) &= \sum_{i=0}^{m} C_{i,m}(x + a_1)^{2(m-i)+\sigma}. \end{aligned} \tag{38} \qquad (\sigma = 0, 1; C_{i,m} \text{ constants})$$

Comparing (37, 38), we obtain the important relation

$$\phi_{2m+\sigma}(x) \equiv \sum_{i=0}^{m} C_i \binom{2m+\sigma}{2i}(x + a_1)^{2(m-i)+\sigma} \tag{39}$$

$$(\sigma = 0, 1; m = 0, 1, \cdots; C_i \text{ constants}),$$

where the C_i are functions of the a_i ($i = 1, 2, \cdots$) and do not depend explicitly on m ($C_0 = 1$). We note the additional property of (39) that only even or only odd powers of $x + a_1$ occur. Moreover, by the orthogonality property,

$$\int_{-\infty}^{\infty}(x + a_1)^i\phi_n(x)d\psi(x) = \int_{-\infty}^{\infty}x^i\phi_n(x - a_1)d\psi(x - a_1) = 0 \tag{40}$$

$$(i = 0, 1, \cdots, n - 1).$$

Letting

$$\beta_i \equiv \int_{-\infty}^{\infty}(x + a_1)^i d\psi(x) = \int_{-\infty}^{\infty}x^i d\psi(x - a_1), \quad \beta_0 > 0 \qquad (i = 0, 1, \cdots),$$

we have [11]

$$\phi_n(x - a_1) = \frac{1}{\Delta_n}\begin{vmatrix} \beta_0 & \beta_1 & \cdots & \beta_n \\ \beta_1 & \beta_2 & \cdots & \beta_{n+1} \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ \beta_{n-1} & \beta_n & \cdots & \beta_{2n-1} \\ 1 & x & \cdots & x^n \end{vmatrix}, \qquad \Delta_n = \begin{vmatrix} \beta_0 & \cdots & \beta_{n-1} \\ \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots \\ \beta_{n-1} & \cdots & \beta_{2n-2} \end{vmatrix}.$$

We see that the sequence $\phi_n(x - a_1)$, $n = 0, 1, \cdots$, is a sequence of *symmetric* orthogonal polynomials whose moments are $\beta_0, \beta_1, \cdots$. If we take $n = 1, 2, \cdots$ successively in (40), it follows that

$$\beta_{2i+1} = 0 \tag{41} \qquad (i = 0, 1, \cdots).$$

Substituting (39) in the first integral (40), we obtain now (for $\sigma = 0, 1$) the $2m$ equations

$$C_m\beta_0 + \binom{2m}{2}C_{m-1}\beta_2 + \cdots + \binom{2m}{2m-2}C_1\beta_{2m-2} + \beta_{2m} = 0,$$

. . . .

$$C_m\beta_{2m-2} + \binom{2m}{2}C_{m-1}\beta_{2m} + \cdots + \binom{2m}{2m-2}C_1\beta_{4m-4} + \beta_{4m-2} = 0,$$

$$(42)\qquad (2m+1)C_m\beta_2 + \binom{2m+1}{3}C_{m-1}\beta_4 + \cdots + \binom{2m+1}{2m-1}C_1\beta_{2m} + \beta_{2m+2} = 0,$$

. . . .

$$(2m+1)C_m\beta_{2m} + \binom{2m+1}{3}C_{m-1}\beta_{2m+2} + \cdots$$
$$+ \binom{2m+1}{2m-1}C_1\beta_{4m-2} + \beta_{4m} = 0.$$

In particular,

$$C_1\beta_0 + \beta_2 = 0.$$

By means of (42), we may eliminate $C_1, C_2, \cdots, C_m$ and express β_{2i} ($i = 2, 3, \cdots$) uniquely in terms of β_0, β_2. Evidently, we may assume β_0 fixed (say $= 1$). Hence, β_{2i}, C_i ($i = 1, 2, \cdots$) are determined as soon as β_2 (> 0) is given.

It is clearly sufficient to show that all systems of polynomials given in the form (39), which satisfy (40), are reducible to one system. Let $\{\bar\phi_n(x)\}$ be such a system where a_1, β_2 have certain given values $\bar a_1, \bar\beta_2$ respectively. By the preceding, β_{2i}, C_i for the system $\{\bar\phi_n(x)\}$ will have certain values $\bar\beta_{2i}, \bar C_i$ ($i = 1, 2, \cdots$). Consider now another such system $\{\phi_n(x)\}$ determined by different a_1, β_2. Set $\beta_2 = c^2\bar\beta_2$ ($c =$ certain positive constant). Then, by induction from (42), we get for the β_{2i}, C_i corresponding to the system $\{\phi_n(x)\}$, the important relations

$$\beta_{2i} = c^{2i}\bar\beta_{2i}, \qquad C_i = c^{2i}\bar C_i \qquad (i = 1, 2, \cdots).$$

It follows by (39) that

$$\begin{aligned}\phi_{2m+\sigma}(x + \bar a_1 - a_1) &= \sum_{i=0}^{m} C_i\binom{2m+\sigma}{2i}(x + \bar a_1 - a_1 + a_1)^{2(m-i)+\sigma}\\ &= \sum_{i=0}^{m} c^{2i}\bar C_i\binom{2m+\sigma}{2i}(x+\bar a_1)^{2(m-i)+\sigma}\\ &= c^{2m+\sigma}\sum_{i=0}^{m}\bar C_i\binom{2m+\sigma}{2i}\left(\frac{x+\bar a_1}{c}\right)^{2(m-i)+\sigma},\end{aligned}$$

or, again, if we apply (39) to $\bar\phi_{2m+\sigma}(x)$,

$$\phi_{2m+\sigma}(cx + c\bar a_1 - a_1) \equiv c^{2m+\sigma}\bar\phi_{2m+\sigma}(x).$$

Thus all systems of orthogonal, Appell polynomials are reducible, by the linear transformation $x/cx + c\bar{a}_1 - a_1$ to one system, which is necessarily the system of Hermite polynomials

$$H_n(x) = x^n - \frac{n(n-1)}{1!\,2^2}x^{n-2} + \frac{n(n-1)(n-2)(n-3)}{2!\,2^4}x^{n-4} - \cdots,$$

since the latter is a system of Appell polynomials.

Bibliography

1. Th. Stieltjes, *Sur les polynomes de Jacobi*, Comptes rendus, vol. 100 (1885), pp. 620–622.
2. K. Shibata, *On the distribution of the roots of a polynomial satisfying a certain differential equation of the second order*, Jap. J. Math., vol. 1 (1924), pp. 147–153.
3. W. Lawton, *On the zeros of certain polynomials related to Jacobi and Laguerre polynomials*, Bull. Amer. Math. Soc., vol. 38 (1932), pp. 442–448.
4. D. Hilbert, *Ueber die Discriminante der im endlichen abbrechenden hypergeometrischen Reihen*, J. für Math., vol. 103 (1888), pp. 337–345.
5. F. Klein, *Ueber die Nullstellen der hypergeometrischen Reihe*, Math. Annalen, vol. 37 (1890), pp. 573–590.
6. E. Van Vleck, *A determination of the number of real and imaginary roots of the hypergeometric series*, Trans. Amer. Math. Soc., vol. 3 (1902), pp. 110–131.
7. A. Hurwitz, *Über die Nullstellen der hypergeometrischen Funktion*, Math. Annalen, vol. 64 (1907), pp. 517–560.
8. E. R. Neumann, *Beiträge zur Kenntnis der Laguerreschen Polynome*, Jahresber. der Deut. Math. Ver., vol. 30 (1921), pp. 15–35.
9. C. Winston, *On mechanical quadratures formulae involving the classical orthogonal polynomials*, Annals of Math., vol. 35 (1934), pp. 658–677.
10. W. Hahn, *Die Nullstellen der Laguerreschen und Hermiteschen Polynome*, Schriften des Math. Sem. und des Inst. für angewandte Math. der Univ. Berlin, vol. 1 (1933), pp. 213–244.
11. J. Shohat, *Théorie générale des polynomes orthogonaux de Tchebichef*, Mémorial des Sc. Math., vol. 66, Paris, 1934.
12. J. Meixner, *Orthogonale Polynomsysteme mit einer besonderen Gestalt der erzeugenden Funktion*, Jour. London Math. Soc., vol. 9 (1934), pp. 6–13.
13. P. Appell, *Sur une classe de polynomes*, Ann. Sc. Éc. Norm. Sup., vol. 9 (1880), pp. 119–144.
14. C. Buell, *The zeros of Jacobi and related polynomials*, this Journal, vol. 2 (1936), pp. 304–316.
15. J. Shohat, *The relation of the classical orthogonal polynomials to the polynomials of Appell*, Am. Jour. Math., vol. 58 (1936), pp. 453–464.

The University of Pennsylvania.

RINGS OF SETS

By Garrett Birkhoff

1. **Definitions.** Following Hausdorff,[1] a family $\mathfrak{F}$ of subsets of a class I is said to form a "ring" if and only if it contains, with any two sets[2] S and T, their *sum* (or union) $S \cup T$ and their *product* (or intersection) $S \cap T$. Clearly a ring contains, with any finite number of subsets $S_1, \cdots, S_n$, their sum $S_1 \cup \cdots \cup S_n$ and their product $S_1 \cap \cdots \cap S_n$.

The family $\mathfrak{F}$ is said to constitute a "complete ring" if and only if it contains, with any subfamily $\mathfrak{S}$ of sets S_α, their sum $\underset{\alpha\in\mathfrak{S}}{V} S_\alpha$ and their product $\underset{\alpha\in\mathfrak{S}}{\Lambda} S_\alpha$. The family $\mathfrak{F}$ is also said to be a "σ-ring" if and only if it contains, with any *countable* subfamily $\mathfrak{S}$ of sets S_α, their sum $\underset{\alpha\in\mathfrak{S}}{V} S_\alpha$ and their product $\underset{\alpha\in\mathfrak{S}}{\Lambda} S_\alpha$.

It is obvious that rings containing only a finite number of sets, and σ-rings containing only a countable number of sets, are necessarily complete rings. These theorems can be improved by using chain conditions; however, the family $\mathfrak{C}$ of all finite sets of integers is a countable ring which is not a σ-ring (and a fortiori not complete), while the family $\mathfrak{D}$ of all countable subsets of the continuum is a σ-ring which is not complete.

2. **The importance of the subject.** Rings of sets are mathematically important for a number of reasons. They are conceptually important because one can define them so simply in terms of two fundamental operations. They are also important because the sets of any class I carried within themselves by any one-valued transformation of I into itself are a complete ring. (The proof of this will be left to the reader.) Also, as is well known, the open and closed subsets of any topological space constitute rings, and the measurable subsets of any Cartesian n-space constitute a σ-ring.

Again, the reader will immediately see that

(2α) The sets common to all the rings (resp. σ-rings or complete rings) of any aggregate of rings of subsets of any class I themselves form a ring (resp. σ-ring or complete ring).

It follows that the closed subsets of any topological space Σ invariant under any group of transformations constitute a ring. The study of these rings is important in dynamics,[3] where, however, the existence of minimal closed and connected constituents introduces special considerations. It follows also that

Received January 16, 1937.

[1] *Mengenlehre*, 1927 (2d ed.), p. 77.

[2] We shall systematically use small Latin letters to denote elements, Latin capitals to denote sets of elements, and German capitals to denote families of sets.

[3] Especially in the theory of so-called "central motions". Cf. G. D. Birkhoff, *Dynamical Systems*, 1927, Chap. VII, §6 ff.

the subsets of any class I carried within themselves under any aggregate A of one-valued transformations τ_α of I into itself form a complete ring.

Moreover, all complete rings of sets belong to at least one aggregate A of transformations in this way. More precisely, any complete ring $\mathfrak{R}$ of sets belongs to the "groupoid"[4] of all one-valued transformations carrying every $S \in \mathfrak{R}$ into itself. This shows that rings of sets play the same rôle in the theory of groupoids of one-valued transformations as is played by transitivity and intransitivity in the theory of groups of permutations (one-one transformations).[5]

3. **Equivalent notions.** If we add the empty set O and the all-set I to any complete ring of subsets of a class I, we still have a complete ring. Hence the theory of rings of sets is contained in that of rings containing O and I.

THEOREM 1. *The complete rings of subsets of I which contain O and I can be identified with the different quasi-orderings of I or with the different completely distributive topologies on I.*

Explanation 1. By a "quasi-ordering" of I is meant a binary relation $x \geqq y$ satisfying

P1: $x \geqq x$ (reflexiveness),

P2: $x \geqq y$ and $y \geqq z$ imply $x \geqq z$ (transitivity).

By a "completely distributive topology" is meant a unary operation $S \to \bar{S}$ (called closure) on the subsets of I which satisfies

C1: $\bar{S} \geqq S$, C2: $\bar{O} = O$, C3: $\bar{\bar{S}} = \bar{S}$;

C4: if $S = V_\alpha S_\alpha$, then $\bar{S} = V_\alpha \bar{S}_\alpha$.

(These are related to well-known axioms of Hausdorff on "partial ordering", and of Riesz-Kuratowski on closure.)

Explanation 2. By an "identification" we mean a one-one correspondence preserved under all permutations of the elements of I. It follows that if we call two families of sets of I resp. two relations on I resp. two operations in I "equivalent" if and only if there exists a permutation of the elements of I carrying one into the other, then the numbers of non-equivalent rings of sets, of non-equivalent quasi-orderings, and of non-equivalent completely distributive

[4] A family G of one-valued transformations of I into itself is termed a "groupoid" if and only if it contains the identity ι: $x \to x$ and the product $\sigma\tau$: $x \to \tau[\sigma(x)]$ of any two of its members σ and τ. The author is preparing an article on groupoids in collaboration with S. Ulam.

[5] The sets invariant (i.e., the sets identical with, and not merely supersets of, their transforms) under any permutation or set of permutations constitute a "complete field"—i.e., a complete ring which contains, with any set, its complement. Moreover, any complete field belongs to the group of all transformations leaving its subsets invariant ("intransitive" on its subsets)—this leads to the usual partial descriptions of groups of permutations through their "transitive systems".

Actually, in the case of *groups* of permutations of I, any subset carried within itself under all their permutations is necessarily invariant.

topologies on I are the same—as well as the numbers of distinct complete rings of sets, of distinct quasi-orderings, and of distinct completely distributive topologies.

Proof of theorem. Let $\mathfrak{R}$ be any complete ring of sets containing O and I. Make the definitions: (1) $x \geqq y$ ($\mathfrak{R}$) means that every $S \epsilon \mathfrak{R}$ containing x contains y, and (2) $\bar{S}$ is the product of all sets $T_\alpha \epsilon \mathfrak{R}$ containing S. That the relation and operation so introduced satisfy P1–P2 and C1–C2 is obvious; it is also obvious that the correspondence between them and $\mathfrak{R}$ is preserved under all permutations of the elements of I.

To prove C3–C4, recall that, since $\mathfrak{R}$ is a complete ring, $\bar{S}$ is the least set in $\mathfrak{R}$ containing S. This proves C3 and

(3α) $S \epsilon \mathfrak{R}$ if and only if $S = \bar{S}$.

Now suppose $S = \underset{\alpha}{V} S_\alpha$. Clearly $S \geqq S_\alpha$ irrespective of α; hence $\bar{S} \geqq \underset{\alpha}{V} \bar{S}_\alpha$. But conversely $\underset{\alpha}{V} \bar{S}_\alpha \epsilon \mathfrak{R}$, since $\mathfrak{R}$ is a complete ring, and $\underset{\alpha}{V} \bar{S}_\alpha \geqq \underset{\alpha}{V} S_\alpha = S$; hence $\underset{\alpha}{V} \bar{S}_\alpha \geqq \bar{S}$. This proves C4.

It remains to prove that *every* quasi-ordering and *every* completely distributive topology belong to such an $\mathfrak{R}$, and that distinct $\mathfrak{R}$ determine distinct quasi-orderings and distinct topologies—four assertions in all.

By (3α), if $\mathfrak{R} \neq \mathfrak{R}'$, then certainly $\mathfrak{R}$ and $\mathfrak{R}'$ yield distinct topologies. This proves one assertion. We next wish to prove that

(3β) *Every* completely distributive topology is determined by a suitable $\mathfrak{R}$.

Under any such topology, consider the family $\mathfrak{F}$ of "closed" sets $S = \bar{S}$. Clearly $\mathfrak{F}$ contains O, I, and (by C4) $\underset{\alpha}{V} S_\alpha$ if it contains every S_α. But it also contains $\underset{\alpha}{\Lambda} S_\alpha$ under the same hypotheses.

Proof. If $\bar{S}_\alpha = S_\alpha$ for every α, then $\underset{\alpha}{\Lambda} \bar{S}_\alpha = \underset{\alpha}{\Lambda} S_\alpha$, and so $\overline{(\underset{\alpha}{\Lambda} S_\alpha)} \leqq S_\alpha$ for all S_α, whence $\overline{(\underset{\alpha}{\Lambda} S_\alpha)} \leqq \underset{\alpha}{\Lambda} \bar{S}_\alpha = \underset{\alpha}{\Lambda} S_\alpha$, and so by C1 $\underset{\alpha}{\Lambda} S_\alpha$ is closed. Hence $\mathfrak{F}$ is a complete ring of sets with O and I. Moreover, if S is any set, then $\bar{S}$ is the product of the $T_\alpha \epsilon \mathfrak{F}$ containing S—by Cl, $\underset{\alpha}{\Lambda} T_\alpha = \underset{\alpha}{\Lambda} \bar{T}_\alpha \geqq \bar{S}$, and, by C1–C3, $\bar{S}$ is a closed set $T_\alpha \geqq S$. Thus $\mathfrak{F}$ "determines" the given topology. This proves (3β).

Again, if $\mathfrak{R}$ is given, then

(3γ) $S \epsilon \mathfrak{R}$ if and only if $x \epsilon S$ and $x \geqq y$ ($\mathfrak{R}$) imply $y \epsilon S$.

Proof. If $S \epsilon \mathfrak{R}$, by definition the second statement holds. Conversely if the second statement holds, then S contains, with every x, the set $S(x)$ of all $y \leqq x$ ($\mathfrak{R}$)—i.e., the product of the $S_\alpha \epsilon \mathfrak{R}$ with $x \epsilon S_\alpha$; obviously $S(x) \epsilon \mathfrak{R}$—and so S is the sum $VS(x)$ of the $S(x)$ of the $x \epsilon S$, and is in $\mathfrak{R}$. By (3γ), if $\mathfrak{R} \neq \mathfrak{R}'$, then $\mathfrak{R}$ and $\mathfrak{R}'$ determine different quasi-orderings.

Finally, every quasi-ordering ρ is determined by some $\mathfrak{R}$. For, given ρ, let $\mathfrak{R}(\rho)$ consist of all S such that $x \epsilon S$ and $x \geqq y$ imply $y \epsilon S$. Clearly $O \epsilon \mathfrak{R}(\rho)$ and $I \epsilon \mathfrak{R}(\rho)$. Also, if a family $\mathfrak{S}$ of S_α is in $\mathfrak{R}(\rho)$, then $(VS_\alpha) \epsilon \mathfrak{R}(\rho)$ and

$(\Lambda S_\alpha) \in \mathfrak{R}(\rho)$. Thus $\mathfrak{R}(\rho)$ is a complete ring of sets containing O and I. Further, if $x \geqq y$, then obviously $x \geqq y$ $(\mathfrak{R}(\rho))$ in the sense that $x \in S \in \mathfrak{R}(\rho)$ and $x \geqq y$ imply $y \in S$. Conversely if $x \geqq y$ $(\mathfrak{R}(\rho))$, then the set $S(x)$ of all $z \leqq x$ (which is in $\mathfrak{R}(\rho)$ by P2 and contains x by P1) contains y—by definition of $x \geqq y$ $(\mathfrak{R}(\rho))$—and so $x \geqq y$. This proves the fourth assertion.

4. **The case of fields of sets.** Which quasi-orderings and which completely distributive topologies correspond to complete fields[6] of sets? And what does this make Theorem 1 reduce to for fields of sets?

THEOREM 2. *In Theorem* 1, *a quasi-ordering corresponds to a* (*complete*) *field of sets if and only if it is an equivalence relation*; *a topology does, if and only if the closures of its points are the subsets of a partition of* I.

Explanation. By an "equivalence relation" is meant a quasi-ordering which satisfies

P3′: $x \geqq y$ implies $y \geqq x$.

By a "partition" of a class I is meant a division of its elements into disjoint subsets, whose sum is I.

Proof. Let $\mathfrak{R}$ be a complete ring of sets, and let $S(x)$ be the product of the sets $S_\alpha \in \mathfrak{R}$ containing x. Then $x \geqq y$ $(\mathfrak{R})$ means $y \in S(x)$. If $\mathfrak{R}$ is a field, and $y \in S(x)$, then the complement $S'(y)$ of $S(y)$ cannot contain x—otherwise $x \in S'(y) \cap S(x) \leqq S(x) - y < S(x)$—and so $y \in S(x)$ implies $x \in S(y)$. This proves P3′. Again, topologically, $S(x)$ is the closure of x. Hence if $\mathfrak{R}$ is a field, unless $S(x)$ and $S(y)$ are disjoint, $S(x) \cap S(y)$ contains some point z, and $x \geqq z$ and $y \geqq z$, whence by P2 and P3′ $x \geqq y$ and $y \geqq x$, and therefore $S(x) = S(y)$.

Conversely, if P3′ holds, and $\mathfrak{R}$ is the family of sets S such that $x \in S$ and $y \leqq x$ imply $y \in S$, then $S \in \mathfrak{R}$ implies that x not in S and $y \leqq x$ imply y not in S (otherwise $y \in S$ and $x \leqq y$ by P3′), whence $S' \in \mathfrak{R}$ and $\mathfrak{R}$ is a field. The fact that the sums of the parts of any partition of I are a complete field of sets is obvious.

COROLLARY.[7] *The complete fields of subsets of I which contain O and I can be identified with the different equivalences on I or with the different partitions of I.*

5. **Rings of sets and distributive lattices.** We shall deal below with rings of sets without assuming completeness.

Suppose we consider rings of sets simply as collections of symbols (forgetting that the symbols denote sets of points) related by inclusion, addition and multiplication. Then any ring of sets appears as a "distributive lattice", or system $\mathfrak{R}$ of elements S, T, U satisfying[8]

[6] A (complete) ring of sets is called a (complete) field if and only if it contains the complement of every one of its members.

[7] Part of this result is proved by H. Hasse, *Höhere Algebra*, vol. I, 1933, p. 15, and B. L. van der Waerden, *Moderne Algebra*, p. 14.

[8] Cf. the author's *On the structure of abstract algebras*, Proc. Camb. Phil. Soc., vol. 31

L1: $S \cap S = S$ and $S \cup S = S$.

L2: $S \cap T = T \cap S$ and $S \cup T = T \cup S$.

L3: $(S \cap T) \cap U = S \cap (T \cap U)$ and $(S \cup T) \cup U = S \cup (T \cup U)$.

L4: $S \cap (S \cup T) = S \cup (S \cap T) = S$.

L6: $S \cup (T \cap U) = (S \cup T) \cap (S \cup U)$ and

$$S \cap (T \cup U) = (S \cap T) \cup (S \cap U).$$

Moreover, two rings of sets seem indistinguishable when and only when they are "isomorphic"—i.e., admit a one-one correspondence preserving inclusion, sums and products.[9]

Conversely, every abstractly given distributive lattice is known to be obtainable from at least one ring of sets.[10]

6. **Representation theory for distributive lattices.** It is generally true in representation theories for abstract algebras that one gets the simplest results by considering homomorphic (many-one) as well as isomorphic (one-one or "true") representations.

A full representation theory for Boolean algebras by fields of sets has been developed by Stone,[11] and it is interesting to see the complications which arise in the more general case of distributive lattices. These show that the assumption that complements exist cannot be eliminated in Stone's theory.

First, let $\mathfrak{R}$ be any distributive lattice, and let θ be any congruence relation[12] on $\mathfrak{R}$. Then the elements congruent to O form an "ideal" $\mathfrak{O}$ in the sense that

I1: $X \in \mathfrak{O}$ and $A \in \mathfrak{R}$ imply $A \cap X \in \mathfrak{O}$.

I2: $X \in \mathfrak{O}$ and $Y \in \mathfrak{O}$ imply $X \cup Y \in \mathfrak{O}$.

In case $\mathfrak{R}$ is a Boolean algebra, $\mathfrak{O}$ determines θ, but this is not generally true in distributive lattices.

Proof. With Boolean algebras, S and T are congruent mod θ if and only if $(S \cap T') \cup (S' \cap T) \in \mathfrak{O}$, whereas O is an ideal in the chain of three elements $I > X > O$, determined by two distinct congruence relations.

(1935), pp. 433–454. O. Ore calls distributive lattices "arithmetic structures". Considerable work has been done by Fritz Klein on the decomposition of distributive lattices important in number theory; M. Ward has also given categorical definitions of such systems.

[9] Actually, any one-one correspondence preserving any one of these three preserves all; this is not true of many-one correspondences.

[10] Cf. the author's *On the combination of subalgebras*, Proc. Camb. Phil. Soc., vol. 29 (1933), pp. 441–464, Theorem 25.2.

[11] M. H. Stone, *The theory of representations of Boolean algebras*, Trans. Amer. Math. Soc., vol. 40 (1936), pp. 37–111. By a "representation" of a distributive lattice L, we mean a homomorphism between L and a ring of sets.

[12] I.e., any partition of the elements of $\mathfrak{R}$ determining an abstract homomorphism. This is a basic notion of general abstract algebra, whose detailed definition we shall omit.

Again, let $\mathfrak{R}$ be any distributive lattice, and A any element of $\mathfrak{R}$. The relation $X \cong Y \bmod A$ meaning $X \cup A = Y \cup A$ is a congruence relation.

Proof. That it is an equivalence relation is obvious. Moreover, by L1–L6,

$$(X \cup Y) \cap A = (X \cup A) \cup (Y \cup A),$$

$$(X \cap Y) \cup A = (X \cup A) \cap (Y \cup A);$$

hence the correspondence $X \to X \cup A$ defines a homomorphism of $\mathfrak{R}$ onto a subring of itself.

If $\mathfrak{R}$ is a finite Boolean algebra, there are no other congruence relations on $\mathfrak{R}$; this is not true for finite distributive lattices which are not Boolean algebras (proof omitted).

7. **Prime ideals.** Let us now suppose that R is any distributive lattice, and let us attempt to give a full representation theory for R.

Let θ: $R \to \mathfrak{R}$ be any homomorphism from R to a ring of subsets of a class I. We may classify the points of I into three categories: those contained in every set $X \in \mathfrak{R}$, those contained in no set $X \in \mathfrak{R}$, and the others. The first two categories of points are trivial, and so we can assume that $O \in \mathfrak{R}$ and $I \in \mathfrak{R}$.

Under these circumstances, every $p \in I$ divides the elements of R into two categories: those corresponding to sets including p, and those corresponding to sets excluding p. The second set of elements is an "ideal", while the first is a "dual ideal" D in the sense that

D1: $x \in D$ and $a \in R$ imply $a \cup x \in D$.

D2: $x \in D$ and $y \in D$ imply $x \cap y \in D$.

Hence the representation of R through $\mathfrak{R}$ is characterized to within equivalence[13] by which divisions of R into an ideal and complementary dual ideal occur, and how many times each occurs.

But conversely, by I1–I2 and D1–D2, if one is given any correspondence associating each division π of R into an ideal J and complementary dual ideal D with a cardinal number $n(\pi)$, then this belongs to a representation of R by a ring of sets, and so if we define (with Stone, op. cit.) an ideal to be "prime" if and only if its complement is a dual ideal, we have

THEOREM 3. *The inequivalent representations of a given distributive lattice R as a ring of sets are the different functions whose arguments are the "prime ideals" of R, and whose values are cardinal numbers.*

Remark 1. With Boolean algebras, the number of elements in any prime ideal and its dual are the same. Also, no prime ideal contains any other prime ideal. Neither of these properties is true in distributive lattices not Boolean algebras (e.g., the chain $I > X > O$).

[13] I.e., to within differences between the various points $p \in I$. This is standard terminology.

Remark 2. It is natural to call a representation "irredundant" if and only if no prime ideal appears as a point more than once.

8. The finite case. Only exceptionally are the prime ideals of infinite Boolean algebras known. But in each finite Boolean algebra of order 2^n they are known to be n sublattices of order 2^{n-1}.

We shall go further and determine the prime ideals of all finite distributive lattices.

Accordingly, let R be any finite distributive lattice, P any prime ideal in R, and $D = R - P$ the dual of P. Form any connected chain[14] $O < x_1 < x_2 < \cdots < x_r = I$ in R; it is clear that in such a chain there will be exactly one "link" $x_i < x_{i+1}$ such that $x_i \in P$ and $x_{i+1} \in D$, and that $x_k \in P$ for $k \leqq i$, while $x_k \in D$ for $k > i$.

(8α) We have $y \in P$ or $y \in D$ according as $v = x_i \cup (y \cap x_{i+1}) = (x_i \cup y) \cap x_{i+1}$ is x_i or x_{i+1}.

Proof. Since $x_i \leqq x_{i+1}$, $x_i \cup (y \cap x_{i+1}) = (x_i \cup y) \cap x_{i+1}$ (by L6 and contraction). Again, for any y, obviously $x_i \leqq v \leqq x_{i+1}$; hence either $v = x_i$ or $v = x_{i+1}$ (no further interpolation being possible). But if $[x_i \cup (y \cap x_{i+1})] = x_i \in P$, then by I1 $(y \cap x_{i+1}) \in P$, and so by D2 (since $x_{i+1} \in D$), $y \in P$. Similarly, if $x_i \cup (y \cap x_{i+1}) = x_{i+1} \in D$, then by I2 (since $x_i \in P$), $y \cap x_{i+1} \in D$, and so by D1, $y \in D$.

Definition. By a "prime factor" of a distributive lattice is meant any symbol x/y, where $y < x$ and no element can be interpolated between y and x. A prime factor x/y will be called a "cleavage" for a given prime ideal P if and only if $y \in P$ and $x \in (R - P)$.

(8β) Any prime factor a/b is a cleavage for some prime ideal.

Proof. Let $x \in P$ if and only if $(b \cup x) \cap a = b$; this makes $x \in (R - P)$ if and only if $(b \cup x) \cap a = a$, since for all x, $b \leqq (b \cup x) \cap a = b \cup (x \cap a) \leqq a$, and a/b is prime. Clearly $a \in (R - P)$ and $b \in P$ (by L4). It remains to prove I1–I2 and D1–D2. But I1 and D1 are obvious, since $(b \cup x) \cap a$ is decreased resp. increased by substituting $x \cap y$ resp. $x \cup y$ for x. Moreover, under the hypotheses of I2,

$$\begin{aligned} b &= [b \cup (x \cap a)] \cup [b \cup (y \cap a)] = b \cup [(x \cap a) \cup (y \cap a)] \\ &= b \cup [(x \cup y) \cap (x \cup a) \cap (y \cup a) \cap a] = b \cup [(x \cup y) \cap a]. \end{aligned}$$

This proves I2. The proof of D2 is dual.

THEOREM 4. *Let R be any finite distributive lattice, and let its prime ideals be $P_1, \cdots, P_r$. Then in every connected chain $O < x_1 < \cdots < x_r = I$, each x_{i+1}/x_i is a cleavage for just one P_i—whence $r = n$.*

Proof. By (8α), if $P_i \neq P_j$, they can have no cleavage in common, and by (8β), every prime factor is a cleavage for some P_i.

We have the Jordan-Dedekind theorem[15] on the constancy of the number of

[14] A chain is called "connected" (or dense by Ore) if no further terms can be interpolated in it.

[15] R. Dedekind, *Werke*, vol. II, p. 254.

links in connected chains as one corollary, and using Theorem 3, we have the further

Corollary. *A finite distributive lattice has (to within equivalence) exactly one irredundant isomorphic representation as a ring of sets—and the number of points involved is the number of links in its connected chains.*[16]

9. **The finite case** (continued). Let R denote again any finite distributive lattice, let its prime ideals be $P_1, \cdots, P_n$, and let their duals be $D_1, \cdots, D_n$.

Let further $s_i = s(P_i)$ and $p_i = p(D_i)$ be the sum of the $x \in P_i$ resp. the product of the $x \in D_i$. By I2, $s_i \in P_i$, and by D2, $p_i \in D_i$; hence (cf. I1–D1) $x \in P_i$ means $x \leqq s_i$ and $x \in D_i$ means $x \geqq p_i$.

Now let I denote the partially ordered set[17] of the s_i. Call a subset S of I "closed" if and only if $s_i \in S$ and $s_j \leqq s_i$ imply $s_j \in S$.

(9α) R is isomorphic with the ring of "closed" subsets of I under the correspondence $S \rightleftharpoons \bigwedge_{s_i \in S} s_i$.

Proof. Let S be a "closed" subset of I. Then by I1–I2 and D1–D2, $\bigwedge_{s_i \in S} s_i$ is in the P_i corresponding to these s_i, and no others. But given $x \in R$, the subset of $s_i \geqq x$ is closed, $y = \bigwedge_{s_i \geqq x} s_i$ is in the same P_i as x, and hence $y \cup x$ and $y \cap x$ are, and so by (8β) no prime factor can be inserted between them, and $x = y = \bigwedge_{s_i \geqq x} s_i$. Thus the correspondence $x \rightleftharpoons \bigwedge_{s_i \geqq x} s$ is one-one. But it clearly preserves inclusion, while by Theorem 1 the closed subsets of I are a ring of sets. This completes the proof.

Consequently two finite distributive lattices having isomorphic partially ordered sets of s_i are isomorphic. But the converse is obvious, since the s_i are intrinsically defined. Since, finally, if X is any (abstractly given) partially ordered set, the ring of its closed subsets is a distributive lattice having the "closures" of points of X for s_i, we obtain

Theorem 5.[18] *There is a one-one correspondence between the partially ordered sets of n elements and the distributive lattices whose connected chains are of length n.*

In the notation of a previous article (this volume of this Journal, p. 311), by (9α) this is the correspondence $X \rightleftharpoons B^X$.

Remark 1. The connected components $X_1, \cdots, X_s$ of X correspond to the indecomposable direct factors of

$$B^X = (B^{X_1}) \times \cdots \times (B^{X_s})$$

in the direct decompositions of B^X.

[16] Cf. *On the combination of subalgebras*, Theorem 17.2.

[17] A set is "partially ordered" (the terminology is Hausdorff's, *Grundzüge der Mengenlehre*, 1914, Chap. VI) by a quasi-ordering satisfying P3: $x \geqq y$ and $y \geqq x$ imply $x = y$. Any subset of a partially ordered set (such as a distributive lattice) is partially ordered by the same relation.

[18] Theorem 5 was announced without proof by the author in a note *Sur les espaces discrets*, Comptes Rendus, vol. 201 (1935), p. 19.

Remark 2. The "Hasse[19] diagram" for X gives an infinitely more compact and intelligible way of writing down a general distributive lattice B^X than the multiplication table used by Dedekind, or than the "Hasse diagram" for B^X itself used by recent authors.

Remark 3. Let $\mathfrak{R}$ be any finite ring of sets, and $L = B^X$ the distributive lattice isomorphic with $\mathfrak{R}$. To find X, take the quasi-ordering determined by $\mathfrak{R}$; identify points x and y satisfying both $x \geqq y$ $(\mathfrak{R})$ and $y \geqq x$ $(\mathfrak{R})$; the partial ordering induced by the quasi-ordering on the sets of "identified" points will yield X.

10. **The indecomposable elements.** Let us again suppose that R is a finite distributive lattice with prime ideals $P_1, \cdots, P_n$ having duals $D_1, \cdots, D_n$.

(10α) Each $s(P_i) = s_i$ is product-indecomposable. Dually, each $p(D_i) = p_i$ is sum-indecomposable.

Explanation. An element a of a lattice R is called "product-indecomposable" when no two elements $x > a$ and $y > a$ exist with $x \cap y = a$; it is called sum-indecomposable when (dually) no two elements $x < a$ and $y < a$ exist with $x \cup y = a$.

Proof. If $x > s_i$ and $y > s_i$, then $x \in D_i$ and $y \in D_i$, whence $(x \cap y) \in D_i$ and $x \cap y \neq s_i$.

(10β) If $x \in R$ is product-indecomposable, it is an s_i. If it is sum-indecomposable, it is a p_i, dually.

Proof. If x is product-indecomposable, then it yields a unique prime factor a/x; let P_i be the corresponding prime ideal. Clearly $x = s(P_i)$, since if $x < s(P_i)$, we would have $y = s(P_i)$, whence $y \in P_i$.

COROLLARY. *The number of sum-indecomposable resp. product-indecomposable elements of a finite distributive lattice is the length of its connected chains.*

10a. **The free distributive lattices.** Consider the "free" distributive lattice generated by n symbols $x_1, \cdots, x_n$; Theorem 1 inclines one to adjoin to it elements O and I such that $O < x_i < I$ for all x_i.

If this is done, then the product-indecomposable elements form a Boolean algebra B^n of 2^n elements. More precisely, they are the elements, O, x_k, $x_i \cup x_j$, $x_i \cup x_j \cup x_k$, $x_h \cup x_i \cup x_j \cup x_k$, $\cdots$, $x_1 \cup \cdots \cup x_{k-1} \cup x_{k+1} \cup \cdots \cup x_n$, $x_1 \cup \cdots \cup x_n$. The corresponding prime factors are

$$x_1 \cap \cdots \cap x_n/0,$$

$$x_k \cup (x_1 \cap \cdots \cap x_{k-1} \cap x_{k+1} \cap \cdots \cap x_n)/x_k,$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$(x_1 \cup \cdots \cup x_{k-1} \cup x_{k+1} \cup \cdots \cup x_n) \cup x_k \Big/ (x_1 \cup \cdots \cup x_{k-1} \cup x_{k+1} \cup \cdots \cup x_n),$$

$$I/(x_1 \cup \cdots \cup x_n).$$

[19] Cf. H. Hasse, *Höhere Algebra*, vol. II, 1927, p. 103, p. 123.

It is a corollary that the free distributive lattices with O and I adjoined are the B^{B^n}.

The proofs of the above statements are tedious; they depend on the knowledge of canonical expressions for the elements of the free distributive lattice.[20]

11. **A general decomposition theorem.** Let R be any (finite or infinite) distributive lattice, and let

$$x = x_1 \cap \cdots \cap x_r = y_1 \cap \cdots \cap y_s$$

be any two representations of an element $x \in R$ as a product. Then irrespective of i, $x_i = x_i \cup x = x_i \cup (\Lambda_j y_j) = \Lambda_j (x_i \cup y_j)$. Hence if x_i is product-indecomposable, some $x_i \cup y_j = x_i$. This means some $y_j \leqq x_i$. Symmetrically, some $x_k \leqq y_j$. Hence either $x_k = y_j = x_i$, or x_i is *redundant* in the strong sense that some $x_k < x_i$, whence

$$x = x_1 \cap \cdots \cap x_{i-1} \cap x_{i+1} \cap \cdots \cap x_r .$$

Thus if the decompositions are irredundant, the x_i and y_j are equal in pairs, $r = s$, and so

(11α) *In a distributive lattice, no element has more than one irredundant product-decomposition (sum-decomposition) into elements not themselves further decomposable.*

But conversely, any modular lattice which is not distributive is known[21] to contain a sublattice of five elements a, b, x_1, x_2, x_3 satisfying $a < x_i < b$, $x_i \cap x_j = a$, and $x_i \cup x_j = b$ $[i \neq j]$. Now starting with the two product-decompositions $a = x_1 \cap x_2$ and $a = x_2 \cap x_3$ of a, making further decompositions as long as possible, and eliminating redundant components, we see that any factor for the second decomposition which contains x_1 must contain x_2 or x_3 and hence b—whereas the first decomposition and those derived from it must possess at least one factor containing x_1 but *not* b. Hence if the above process is terminating, we will surely get two distinct product-decompositions of a. But in the presence of the chain-condition, the process *is* terminating. This completes the proof of

THEOREM 6.[22] *A modular lattice satisfying the ascending chain condition is distributive if and only if each of its elements has a unique irredundant product-decomposition.*

[20] The latter are given by Th. Skolem, *Über gewisse "Verbände" oder "Lattices"*, Avh. Norske Videnskaps Akademi i Oslo, Mat.-Naturv. Klasse, 1936, no. 7, pp. 7, 8. From them it is immediately obvious that the elements specified above are the *only* product-indecomposable elements—but there are just enough such elements to give the lattice 2^n dimensions; hence they are *all* product-indecomposable.

[21] Cf. Theorem 4 of the author's *On the lattice theory of ideals*, Bull. Amer. Math. Soc., vol. 40 (1934), p. 617.

[22] This result was announced by the author in Abstract 41-1-75 of the Bull. Amer. Math Soc., vol. 41 (1935), p. 32.

This result is especially interesting in the light of recent proofs by Kurosch and Ore that, in any modular lattice, the number of factors in any two irredundant product-decompositions of the same element into indecomposable factors is the same.[23]

12. **Some enumeration problems.** One very impartial test of one's ability to classify finite systems is one's ability to enumerate them. This suggests the problem of determining the following combinatorial functions.

(12.1) The number $F_1(n)$ of different rings of subsets of n elements. (This is the number of sublattices of the Boolean algebra B^n of 2^n elements.)

(12.2) The number $F_2(n)$ of non-equivalent rings of such subsets. (This is the number of such sublattices non-conjugate under the group of automorphisms of B^n.)

(12.3) The number $F_3(n)$ of non-isomorphic rings of such subsets. (This is the number of non-isomorphic distributive lattices of "dimensions" n.)

(12.4) The number $F_4(n)$ of non-isomorphic partial orderings of n elements.

Remark 1. If we replace "ring" by "field" in the above, $F_1(n)$ becomes a known combinatorial function defined by the recurrence

$$H^* (n + 1) = \sum_{h=0}^{n} \binom{n}{h} H^* (n - h).$$

This has been studied by Aitken (Edin. Math. Notes, vol. 28 (1933), pp. xviii–xxiii). Again, $F_2(n)$ becomes the partition function—a celebrated asymptotic formula for which has been given by Hardy and Ramanujan. And lastly, $F_3(n)$ becomes n.

Remark 2. In virtue of Theorems 3 and 5, $F_3(n) = \sum_{k=1}^{n} F_4(n)$. Also, $F_2(n)$ is by Theorem 1 the number of non-equivalent quasi-orderings of n elements, and $F_1(n)$ is the number of *different* quasi-orderings of n elements.

A table for these functions for small n follows.

	1	2	3	4	5	6
$F_1(n)$	1	3	29			
$F_2(n)$	1	3	9	30		
$F_4(n)$	1	2	5	15	51	250

[23] A. Kurosch, *Durchschnittsdarstellungen mit irreduziblen Komponenten in Ringen und in sog. Dualgruppen*, Rec. Math. (Moscow), vol. 42 (1935), pp. 613–16. O. Ore, *On the foundations of abstract algebra.* II, Annals of Math., vol. 37 (1936), p. 270, Theorem 11. The result of Kurosch-Ore contains a decomposition theorem of E. Noether for ideals, and a less well-known result of Remak's on finite groups, as special corollaries.

In calculating these values, assume that $F_2(n)$ is the number of functions from the different partially ordered sets of $k \leqq n$ elements to cardinal numbers whose sum is n. Also $F_1(n)$ can be calculated combinatorially from $F_2(n)$ by summing the occurrences of each type of ring of subsets, over the types existing. To find $F_4(n)$, separate each partially ordered set into its connected components.

It would be very interesting to know more about the $F_k(n)$, numerically or asymptotically. $F_4(n)$ resembles the function describing the number of groups of order 2^n—whose first values are 1, 2, 5, 14, 51, 266, $\cdots$. It appears to increase more rapidly than the function describing the number of non-isomorphic symmetric relations between n objects (or alternatively, the number of non-homomorphic graphs with n vertices), whose first values are 1, 2, 4, 11, 27. But as almost nothing is known about the rate of growth of these functions, these comparisons are not very reliable.

13. **Homomorphic images and sublattices.** Let us try to determine the homomorphisms and sublattices of a given finite distributive lattice, guided by the previous results.

Some authors,[24] inspired by the numerous analogies between lattices and rings, have correlated the congruence relations on lattices with "ideals" and "normal sublattices". But except in the "complemented" case in which each element x possesses a complement x' with $x \cap x' = 0$ and $x \cup x' = I$, this correlation is incomplete.

Actually, in the case of finite distributive lattices, and more generally with arbitrary modular lattices of finite dimensions (the author will publish proofs elsewhere, in an article on modular lattices), congruence relations correspond one-one to subsets of the set of prime factors.

It follows that, if $L = B^X$ is any finite distributive lattice, the congruence relations on it are obtained by setting $x \equiv y(\theta)$ if and only if x and y contain the same P_i. Hence, to obtain the homomorphic images B^Y of B^X, set Y equal to any subset of X having on that subset the same inclusion relation as X.

The determination of the sublattices of B^X is even easier. First, recall that B^X is isomorphic with the ring of subsets of X which are "closed" with respect to the partial ordering of X, by (9α). But a sublattice is clearly just a subring—and by Theorem 1 these subrings are the families of sets "closed" under quasi-orderings ρ of X such that $x \geqq y(\rho)$ whenever $x \geqq y$ in X. Hence, to obtain the sublattices of B^X, strengthen the inclusion relation in X to any quasi-ordering and consider the partially ordered set Y obtained from this after elements x and y such that $x \geqq y$ and $y \geqq x$ have been identified; B^Y will be the general sublattice of B^X.

HARVARD UNIVERSITY.

[24] Cf. for instance Gr. C. Moisil, *Recherches sur l'algèbre de la logique*, Annales Sci. de l'Univ. de Jassy, vol. 22 (1936), pp. 1–118.

A REPRESENTATION OF GENERALIZED BOOLEAN RINGS

BY N. H. McCOY AND DEANE MONTGOMERY

1. **Introduction.** Stone has recently shown[1] that every Boolean ring is isomorphic to a ring of subclasses of some class. As Stone himself remarks, there is a close relation between the representation of Boolean rings and the theory of direct sums of rings. The theorem just stated is clearly equivalent to the theorem that every Boolean ring is isomorphic to a subring of a direct sum of rings F_2.[2] We present here a simple direct proof of this theorem in a somewhat more general case.

A commutative ring R_p is said to be a *generalized Boolean ring of index p* (often abbreviated *p-ring*) if p is a prime and if for every a in R_p it is true that $a^p = a$ and $pa = 0$. A Boolean ring as defined by Stone is therefore a 2-ring.[3] We show here that a p-ring is isomorphic to a subring of a direct sum of rings F_p. The interest of this theorem lies partly in its generality and partly in the simplicity of the proof, which is based on an exploitation of a device used by Alexander and by Alexander and Zippin.[4] Inasmuch as our proof, like Stone's, demands the existence of certain homomorphisms and inasmuch as we prove the existence of these homomorphisms by a method analogous to Stone's,[5] our proof makes use of transfinite induction.

2. **Subrings of direct sums.** For the theorem given here on subrings of direct sums neither of the rings considered need be commutative.

THEOREM 1. *A necessary and sufficient condition that a ring R be isomorphic to a subring of a direct sum of rings K is that for every $a \neq 0$ in R there is a homomorphism h of R into a subring of K such that $h(a) \neq 0$.*

Consider first the necessity of the condition. Assume then that the elements of R are functions f defined on a certain set M with values in K.[6] If f_1 in R is not zero, there is some element m such that $f_1(m) \neq 0$. We obtain a homomorphism of R into a subring of K by making correspond to any f in R the value $f(m)$. This homomorphism is not zero on f_1 and therefore satisfies the condition of the theorem.

Received January 26, 1937.

[1] M. H. Stone, *The theory of representations for Boolean algebras*, Transactions of the American Mathematical Society, vol. 40 (1936), pp. 37–111. See also Garrett Birkhoff, *On the combination of subalgebras*, Proc. Camb. Phil. Soc., vol. 29 (1933), pp. 441–464.

[2] In general, for any prime p, F_p denotes the field of integers reduced modulo p.

[3] When $p = 2$, the commutativity and the fact that $pa = 0$ follow from the assumption $a^p = a$.

[4] Annals of Mathematics, vol. 35 (1934), pp. 389–395; vol. 36 (1935), pp. 71–85.

[5] Loc. cit., pp. 102–104.

[6] A direct sum consists of the set of all such functions.

Turning now to the sufficiency of the condition, let H denote the set of all homomorphisms h of R into any subring of K. Corresponding to each element a of R we define on H the function y_a, with values in K, as follows:

$$y_a(h) = h(a). \tag{1}$$

Since h is a homomorphism, it follows at once that

$$y_{a+b}(h) = h(a + b) = h(a) + h(b) = y_a(h) + y_b(h), \tag{2}$$

and

$$y_{ab}(h) = h(ab) = h(a)h(b) = y_a(h)y_b(h). \tag{3}$$

Thus the correspondence $a \to y_a$ is a homomorphism of R into the ring S of functions y_a, $a \subset R$. To prove that this is actually an isomorphism, we need to show that the function y_a vanishes identically on H only if $a = 0$. This follows almost at once, for we have assumed that if $a \neq 0$ there is an h_1 in H such that $h_1(a) = y_a(h_1) \neq 0$. Thus R is isomorphic to S, and the proof is completed by noting that the ring S is a subring of the ring of all functions on H to K, and is therefore a subring of a direct sum of rings K.

A theorem entirely analogous to Theorem 1 holds for groups or, more generally, for abstract algebras as defined by Garrett Birkhoff.

3. **Imbedding theorem.** The following theorem is demonstrated for any prime in the same way in which Stone demonstrates it for the case $p = 2$.

THEOREM 2. *A p-ring R_p may be imbedded in a p-ring R_p^* which contains a unit element.*

Let us denote the elements of F_p by $\bar{0}, \bar{1}, \cdots, \overline{p-1}$. The elements of R_p^* will be the pairs $(r, \bar{n})$, where r is in R_p and $\bar{n}$ is in F_p. If $(r_1, \bar{n}_1)$ and $(r_2, \bar{n}_2)$ are two pairs of the kind described, their sum is defined to be $(r_1 + r_2, \bar{n}_1 + \bar{n}_2)$ and their product is defined to be $(r_1r_2 + n_2r_1 + n_1r_2, \bar{n}_1\bar{n}_2)$. These elements form a commutative ring under this definition. Clearly p times any element is zero and it can also be verified that any element raised to the p-th power is itself. The ring thus formed is the desired ring, for the elements of the form $(r, \bar{0})$ form a subring isomorphic to R_p and the element $(0, \bar{1})$ is a unit element.

4. **Finite p-rings.** Let F be a p-ring containing a finite number of elements. Since every element a of F satisfies the equation $a^p = a$, it follows at once that if $a^n = 0$ for some positive integer n, then $a = 0$. Thus F does not contain a radical and is therefore known to have a unit element and to be a direct sum of fields.[7] These fields, being subrings of F, are clearly also p-rings.

It will now be shown that F_p is the only field which is a p-ring. Suppose S is such a field, the unit element of S being denoted by e. Then S contains a field F'_p isomorphic to F_p and consisting of the integral multiples of e. Since

[7] B. L. van der Waerden, *Moderne Algebra*, vol. 2, p. 163.

$$x^p - x \equiv x(x - 1) \cdots [x - (p - 1)] \pmod{p},$$

it follows that each element a of S satisfies the equation

$$x^p - x = x(x - e) \cdots [x - (p - 1)e] = 0.$$

Now it is easy to show that a satisfies a unique equation $f(x) = 0$ of minimum degree, with coefficients in F'_p and with leading coefficient unity. Further, $f(x)$ must be irreducible in F'_p, as otherwise S would contain divisors of zero. But $f(x)$ divides $x^p - x$ and therefore is merely one of its linear factors, and thus a is an element of F'_p. These results yield the following theorem:

THEOREM 3. *Every finite p-ring contains a unit element and is a direct sum of fields* F_p.

The following remarks will be useful in proving the existence of homomorphisms in the next section. Let R_p be any p-ring containing a unit element e and let a be any element in R_p. The ring $\{a, e\}$ generated by a and e consists of all polynomials in a and e. Since $a^p = a$ and since $pa = 0$, this ring is finite and since it is a p-ring, it is expressible as a direct sum of fields F_p. Thus there exists a set of non-zero elements $e_1, e_2, \cdots, e_r$ of $\{a, e\}$ with the following properties:

$$e = e_1 + e_2 + \cdots + e_r, \qquad e_i^2 = e_i, \qquad e_ie_j = 0 \qquad (i \neq j). \tag{4}$$

Every element of $\{a, e\}$ is expressible as a linear combination of the elements e_i with coefficients in F_p. Furthermore the elements e_i are linearly independent over F_p. We shall call this set a *basis* of $\{a, e\}$.

5. **Existence of homomorphisms.** Let R_p be an arbitrary p-ring containing a unit element e and let S be a subring of R_p which contains e. If a is an element of R_p not in S, denote by $S(a)$ the subring of R_p generated by S and a. The elements of the ring $S(a)$ are expressible as polynomials in a having coefficients in S with degree at most $p - 1$. Now let e_i $(i = 1, 2, \cdots, r)$ be a basis of $\{a, e\}$ as in the preceding section. Each integral power of a is a linear combination of the e_i's with coefficients in F_p, and since e is also such a linear combination each element b of $S(a)$ may be written in the form

$$b = b_1e_1 + b_2e_2 + \cdots + b_re_r, \tag{5}$$

the coefficients b_i being elements of S. If $c = c_1e_1 + \cdots + c_re_r$ is another element of $S(a)$, it follows from (4) that

$$\begin{aligned} b + c &= (b_1 + c_1)e_1 + \cdots + (b_r + c_r)e_r, \\ bc &= (b_1c_1)e_1 + \cdots + (b_rc_r)e_r. \end{aligned} \tag{6}$$

If $b = 0$, it follows from (5) and (4) that $0 = e_ib = b_ie_i$, and thus

$$b_1b_2 \cdots b_r = b_1b_2 \cdots b_r(e_1 + \cdots + e_r) = 0. \tag{7}$$

We shall now prove the following lemma.

LEMMA. *Let S be a subring of R_p containing the unit e of R_p, and let h be a given homomorphism $S \to F_p$. Then there exists a homomorphism h', $S(a) \to F_p$, such that under h' the images of elements of S are identical with their images under h.*

The symbol P_r will be used to represent the direct sum of r rings S, the elements of P_r being denoted by $(b_1, b_2, \cdots, b_r)$, where each b_i is an element of S. In like manner C_r will be used to represent the direct sum of r rings F_p. Let K denote the ideal in P_r consisting of those elements $(b_1, b_2, \cdots, b_r)$ such that $b_1e_1 + \cdots + b_re_r = 0$. That K is an ideal follows from (6).

Now h induces a homomorphism $(b_1, b_2, \cdots, b_r) \to (b_1^*, b_2^*, \cdots, b_r^*)$ from P_r to C_r, where $b_i \to b_i^*$ by h. Denote by L the ideal in C_r which is the image of K under the induced homomorphism. The ideal L can not contain $(1, 1, \cdots, 1)$, for if $(b_1, b_2, \cdots, b_r) \to (1, 1, \cdots, 1)$, then $b_1b_2 \cdots b_r \neq 0$, and from (7), $(b_1, b_2, \cdots, b_r)$ can not be in K. Therefore L does not include all of C_r. Any ideal in C_r is made up of elements $(x_1, x_2, \cdots, x_r)$, where for a certain fixed set of i's, $x_i = 0$, and for the remaining i's, x_i may take any value in F_p. Since L is not identical with C_r, we may assume that L consists of all elements $(0, \cdots, 0, x_k, \cdots, x_r)$, where $k > 1$, and $x_k, \cdots, x_r$ are arbitrary elements of F_p.

We now set up the correspondence

$$b = b_1e_1 + \cdots + b_re_r \to b_1^* \tag{8}$$

and proceed to show that this is the required homomorphism h'. By (6) this will be a homomorphism $S(a) \to F_p$ provided the indicated correspondence is independent of the representation (5) of any given element b of $S(a)$. If b can also be expressed as $c_1e_1 + \cdots + c_re_r$, it follows that $(b_1 - c_1)e_1 + \cdots + (b_r - c_r)e_r = 0$ and therefore $[(b_1 - c_1)^*, \cdots, (b_r - c_r)^*] = (b_1^* - c_1^*, \cdots, b_r^* - c_r^*)$ is an element of L. From the form we have assumed L to have, it follows that $b_1^* - c_1^* = 0$ and hence $b_1^* = c_1^*$. Thus (8) defines a homomorphism $S(a) \to F_p$. If x is any element of S, then from (8) we find

$$x = xe = x(e_1 + \cdots + e_r) \to x^*$$

and the homomorphism h' coincides with h on S. This completes the proof.

THEOREM 4. *If R_p is any p-ring containing a unit element e and if a is any non-zero element of R_p, there exists a homomorphism h of R_p into F_p such that $h(a) \neq 0$.*

Consider first the ring $\{a, e\}$. This ring is a direct sum of rings F_p and therefore a homomorphism h may be defined on it to F_p in such a way that $h(a) \neq 0$. It remains only to extend h to the remainder of R_p. If the elements of R_p not in $\{a, e\}$ are well ordered, the desired extension may be made, in view of the lemma just demonstrated, by transfinite induction.[8]

[8] Cf. Stone, loc. cit., pp. 102–104.

6. **Representation of p-rings.** We are now in a position to prove our principal theorem.

THEOREM 5. *If R_p is any p-ring, it is isomorphic to a subring of a direct sum of rings F_p.*

By Theorem 2, R_p may be imbedded in a p-ring R'_p with a unit element, and by Theorems 1 and 4, R'_p is isomorphic to a subring of a direct sum of rings F_p. Therefore R_p itself is isomorphic to a subring of a direct sum of rings F_p.

SMITH COLLEGE.

A STRUCTURAL CHARACTERIZATION OF PLANAR COMBINATORIAL GRAPHS

By Saunders Mac Lane

1. **Introduction.** There are several known necessary and sufficient conditions that a combinatorial graph be planar.[1] This paper aims to establish another such condition which has a more intrinsic character, in that it is obtained directly from an analysis of the structure of the graph. More explicitly, the new condition depends on a unique decomposition of the graph into certain maximal triply connected subgraphs. This decomposition can be viewed on its own merits as a generalization of the Whyburn cyclic element theory.

The first combinatorial criterion for a planar graph is due to Kuratowski,[2] who showed that a graph is planar if and only if it contains no subgraph homeomorphic to one of the two following graphs: the graph composed of five vertices and ten edges, in which each pair of vertices is joined by an edge; the graph composed of six vertices, arranged in two sets of three vertices each, and of nine edges, such that each vertex of the first set is joined to each vertex of the second set by an edge. Subsequently, Whitney defined combinatorially the relation between a graph and its planar dual and showed that a graph is planar if and only if it has a planar dual.[3] A third condition states that a combinatorial graph is planar if and only if it contains a complete independent set of circuits, modulo 2, such that no edge appears in more than two of these circuits.[4]

The application of any of these theorems to a particular case has a haphazard character because one must investigate any possible smallest non-planar subgraph or any possible dual or any possible complete set of circuits. We seek an intrinsic condition;[5] that is, a condition expressible in terms of configurations which are associated in a unique manner with a given graph. An example of such a condition is the result of Whitney, that any graph G can be broken up uniquely into non-separable components[6] and that the graph is planar if and only if each of its non-separable components is planar.[7]

Received January 28, 1937; presented to the American Mathematical Society, December 30, 1936.

[1] For definitions of terms see §2.

[2] K. Kuratowski, *Sur le problème des courbes gauches en topologie*, Fundamenta Mathematicae, vol. 15 (1930), pp. 271–283.

[3] H. Whitney, *Non-separable and planar graphs*, Transactions of the American Mathematical Society, vol. 34 (1932), pp. 339–362.

[4] S. Mac Lane, *A combinatorial condition for planar graphs*, Fundamenta Mathematicae, vol. 28 (1937), pp. 22–32.

[5] It is to be hoped that such a condition may throw light on the question of when a graph is mappable on the torus.

[6] These components of G are precisely the true cyclic elements of G. See G. T. Whyburn, *Concerning the structure of a continuous curve*, American Journal of Mathematics, vol. 50 (1928), pp. 167–194. For the combinatorial decomposition into components see D. König, *Theorie der endlichen und unendlichen Graphen*, Leipzig, 1936, Ch. 14.

[7] H. Whitney, loc. cit., Theorems 12 and 27.

These non-separable components are cyclically connected, since they are not disconnected by the removal of any one vertex. Similarly, a graph G is triply connected if the removal of any two of its vertices either does not disconnect G, or else disconnects G into two parts, one of which is only a chain. Some results of Adkisson and Whitney[3] indicate that such a triply connected graph G can have but one map on the sphere. In this topologically unique map, the circuits which bound the connected regions of the complementary set are the only circuits which do not "cut" the graph; that is, they are the only circuits whose removal does not disconnect the graph. The condition that these circuits really give possible region boundaries for a map yields (§4) an intrinsic condition that such a triply connected graph be planar.

It is then natural to try to reduce a graph G which is not triply connected to triply connected constituents. This might be done by choosing two vertices p and q whose removal disconnects the graph G into two subgraphs H and H'. However, planar maps of H and H' cannot always be combined to form a map of G. This would be the case if H is modified by adding one new arc joining p to q. The so-modified graph we call a "block" of G. Any graph can be broken up successively ("split") into blocks so that the final blocks, called atoms, are triply connected (cf. §2). These atoms are uniquely determined (see §3) except for a homeomorphism. In terms of this combinatorial decomposition, we obtain our fundamental result that a non-separable graph is planar if and only if each of its triply connected atoms is planar (see §5). In the last section we further illustrate the applicability of these atoms by showing that the number of topologically distinct maps of a planar graph can be directly computed from the number of atoms and the number of "multiple" splits of the original graph.

2. The splitting process. We first state some preliminary definitions. A *combinatorial graph* G consists of a finite number of elements $\alpha, \beta, \cdots$ called "edges" and a finite set of "vertices", $p, q, \cdots$, where each edge β is "on" exactly two distinct vertices p and q, which may be called the "ends" of β. Any set of edges in G together with all the vertices on these edges form themselves a *subgraph*. We write $H \subset G$ for "H is a subgraph of G". If H_1 and H_2 are subgraphs of G, then $H_1 \cap H_2$ is the subgraph containing those edges in both H_1 and H_2, while $H_1 + H_2$ is the subgraph containing those edges in either H_1 or H_2, and $G - H_1$ is the subgraph containing all edges of G not in H_1. "Circuits" and "chains" are defined as usual. A *hanging chain* in G is a chain none of whose vertices, except perhaps its two ends, are on more than two edges of G. Two graphs G and G' are called *homeomorphic* if and only if G can be changed into G' by one or more of the following operations:

[3] V. W. Adkisson, *Cyclicly connected continuous curves whose complementary domain boundaries are homeomorphic, preserving branch points*, Comptes Rendus des séances de la Société des Sciences et des Lettres de Varsovie, Classe III, vol. 23, pp. 164–193.

H. Whitney, *Congruent graphs and the connectivity of graphs*, American Journal of Mathematics, vol. 54 (1932), pp. 150–168.

(i) The replacement of a hanging chain by a new hanging chain having the same ends;

(ii) The renaming of edges or vertices or both.[9]

If no vertices are renamed in the homeomorphism, each branch vertex of G is a branch vertex of G', and so we say that the *branch points are preserved.* Two homeomorphic graphs are topologically equivalent, in that any map of the one is topologically homeomorphic to any map of the other.

A graph B consisting only of $t \geqq 3$ edges all having the same two ends p and q, or any graph obtained from B by replacing these edges by hanging chains, will be called a *branch graph* with t branches and with the *termini* p and q.

Because of Whitney's results on separable graphs, stated in the introduction, we shall consider henceforth only non-separable graphs G. A graph G, not a single edge, is *non-separable* (cyclically connected) if and only if each pair of vertices of G is contained in a circuit of G.[10]

A *semi-split* of G at the vertices h_1 and h_2 is a representation of G as a sum

$$G = H + H', \tag{1}$$

where H and H' are two non-void subgraphs, having in common no arcs and no vertices except the vertices h_1 and h_2. A *split* of G at h_1 and h_2 is a representation (1) which is a semi-split and where neither H nor H' is a chain. Since G is non-separable, H must contain both h_1 and h_2, and must be connected. Hence there is a chain in H with ends h_1 and h_2, and in like manner, a chain in H' with the same ends.

Take any chains X and X' contained in H' and H respectively, and having the ends h_1 and h_2. The two subgraphs

$$H + X, \qquad H' + X'; \qquad\qquad X \subset H', \qquad X' \subset H \tag{2}$$

will be called *blocks*[11] of G corresponding to the split (1). We say that G has been split into these two blocks. These blocks are not uniquely determined, for X may be replaced by other chains from H' with the same ends. However, any such X is a hanging chain in the block $H + X$, so that this block is uniquely determined up to a homeomorphism.

A graph G will be called *triply connected*[12] if it is cyclically connected, if it

[9] H. Whitney, *On the classification of graphs,* American Journal of Mathematics, vol. 55 (1933), pp. 236–244.

[10] Whitney, *Non-separable graphs,* loc. cit., Theorem 7. Alternatively, a graph is non-separable if it is not disconnected by the removal of any one vertex.

[11] Since $H + X$ is obtained by replacing all arcs of H' by a single chain X, we may consider this block $H + X$ as a sort of "factor-graph" of G modulo H'. This analogy with factor-groups is suggested because the group of cycles of $H + X$ (mod 2) is isomorphic to the factor-group of the group of cycles of G modulo the group of cycles of H', both taken mod 2.

[12] This notion of triply connected (call it TCM) does not always agree with the notion of triply connected (TCW) introduced by Whitney, in *Congruent graphs,* loc. cit., p. 158. The two definitions agree for graphs G containing at least four vertices and containing

has no split, and if G is neither a circuit nor a branch graph. A cyclically connected graph G split into two blocks may be further decomposed by splitting one of these blocks which may happen not to be triply connected. Such successive splits finally yield a set of unsplittable blocks. By the definition of triple connectivity, each unsplittable block is either a branch graph or a triply connected graph. The branch graphs are unimportant. The triply connected blocks we call *atoms,* and all the triply connected blocks in a final set of unsplittable blocks will be said to constitute a *complete set of atoms* of G.

We now list two useful consequences of our definitions.

LEMMA 1. *If in the graph G there is a chain L which does not pass through*[13] *either vertex h_1 or h_2, then L is contained in one of the blocks of any split at the vertices h_1 and h_2.*

The proof follows at once from the definition of a split.

LEMMA 2. *A block of a cyclically connected graph is always cyclically connected.*

Proof. By hypothesis any two vertices p and q in the block $H + X$ of (2) lie together on a circuit C in G. If C does not lie entirely within the block $H + X$, then one piece of C between h_1 and h_2 can be replaced by the chain X, giving a new circuit C^* in the block and connecting the vertices p and q.

3. **A unique characterization of atoms.** This section will give a combinatorial proof that the atoms of a graph are unique, up to a homeomorphism. This will be done by giving an invariant characterization of these atoms as maximal triply connected subgraphs of G. Here a subgraph T of G is *maximal triply connected* (max. trip. conn.) if T is contained in no other triply connected subgraph $W \neq T$.

THEOREM 1. *Every atom of G is a maximal triply connected subgraph. If*

$$A_1, A_2, \cdots, A_m \tag{3}$$

is a complete set of atoms of G, then every maximal triply connected subgraph of G is homeomorphic, preserving branch points, to one and only one of the atoms (3).

Proof. If T is any trip. conn. subgraph of G, then in the split (1), one of the blocks contains a trip. conn. subgraph T^* homeomorphic, preserving branch points, to T. To show this, use the equation

$$T = (T \cap H) + (T \cap H'), \tag{4}$$

where $T \cap H$ and $T \cap H'$ have in common no arcs and only the vertices h_1 and h_2. As T is trip. conn., this cannot be a split, and one of the subgraphs, say $T \cap H'$, is void or a single chain. If it is void, then $T \subset H$, so that T itself is

neither a circuit E, consisting of only two edges, nor a vertex p, on only two edges. A graph containing a circuit E is never TCM, but may be TCW. A graph containing a vertex p, as above, is never TCW, but may be TCM. The definition TCM used above has the advantage that it is invariant under any homeomorphism of the graph.

[13] A chain L does not pass through h if h is not in L or is only an endpoint of L.

in one of the blocks. On the other hand, if $T \cap H'$ is a single chain Y, this chain must have the ends h_1 and h_2, so that

$$T^* = (T - Y) + X \qquad (Y = T \cap H') \tag{5}$$

is a new subgraph homeomorphic to T, preserving branch points, and T^* is contained in the block $H + X$, as required.

To show that any trip. conn. atom A is max. trip. conn., suppose instead that $A \subset T$, $A \neq T$ holds for a trip. conn. T with more edges than A. If we make one of the splits (1), leading up to the construction of the atom A, then A belongs to one of the blocks (2), say to the block $H + X$. But $T \supset A$, so that T contains at least a circuit of H, and the T^* constructed above from T must certainly be in the same block $H + X$ with A. Because $T \supset A$, and because T^* is obtained from T by changing at most one chain not in A, we must have $T^* \supset A$. Therefore, the atom A is contained in a larger trip. conn. subgraph T^* of the block $H + X$. We repeat this argument, getting A in successively smaller blocks until A is itself a block contained in a larger trip. conn. subgraph T_k, which in turn is contained in the block A. This is a manifest impossibility. Hence every atom is max. trip. conn.

Consider any max. trip. conn. subgraph T. We know that it is homeomorphic, preserving branch points, to a T^* in the block $H + X$. This T^* is max. trip. conn. in this block. This is obvious if $T^* = T$. Otherwise $T \neq T^*$, and we have $Y \neq X$ in the construction (5) of T^*. Were T^* contained in a larger trip. conn. subgraph W in this block, then W, perhaps modified by replacing the chain X by the chain Y from (5), would be a trip. conn. subgraph properly containing T, contrary to the assumption that T is maximal. As a result T^* is a max. trip. conn. subgraph homeomorphic to T and contained in the block $H + X$. Continuing this, we finally obtain a max. trip. conn. subgraph T_k homeomorphic, preserving branch points, to the original T and contained in a smallest block A. As this A contains the triply connected T_k, it cannot be a branch graph, and so must itself be trip. conn. The maximal T_k must be all of A, so that the original T is homeomorphic to the atom $A = T_k$. It can be homeomorphic, preserving branch points, to no other atom, for no two blocks[14] and hence no two atoms can contain the same branch points. The theorem is thus established. It states that a set (3) of atoms contains all max. trip. conn. subgraphs of G, except for certain homeomorphisms. Hence we deduce

THEOREM 2. *If a cyclically connected graph G can be split up in two ways to give two complete sets of atoms, then there exists a one-to-one correspondence between these two sets of atoms, so that corresponding atoms are homeomorphic, preserving branch points.*

This theorem could also be proved by a direct consideration of two given splits by a method similar to that of the Jordan-Hölder theorem; that is, by first showing that any two given original splits of G can be subdivided to give homeomorphic results.

[14] Except in the trivial case when one block is a branch graph.

The characterization of atoms given above can be made independent of the notion of a split by means of an independent description of triply connected graphs, based on the result of Whitney that the chief characteristic of a triply connected graph is the fact that any two of its branch points are the termini of three independent arcs. These independent arcs form a θ-subgraph, where a "θ-graph" means a branch graph with three branches.

THEOREM 3. *A cyclically connected graph G is triply connected if and only if*

(i) *for each pair of branch vertices p and q of G there is a θ-subgraph W of G with termini p and q, and*

(ii) *G contains no circuits passing through less than three branch points.*

Proof. Suppose first that conditions (i) and (ii) hold while G still has a split (1). Since G is not a single chain, and since by (ii) H can not be a branch graph, H must contain a branch vertex different from both h_1 and h_2. Similarly, H' contains a branch vertex q distinct from h_1 and h_2. By (i) there is a θ-graph with termini p and q, and each of the three independent arcs of this θ-graph must pass through one of the two vertices h_1 or h_2 to go from p to q. Thus two of these arcs intersect, a contradiction.

Conversely, suppose that G is trip. conn. and suppose, contrary to (ii), that G has a circuit C with only two branch vertices h_1 and h_2. Then $G = C + (G - C)$ would be a semi-split of G at h_1 and h_2. This semi-split must be a split, because were $G - C$ a single chain, G would be merely a θ-graph, contrary to the definition of triple connectivity.

To prove (i) for a trip. conn. G, first modify the graph by replacing each maximal hanging chain of G by a single edge. The resulting graph contains no vertices not branch vertices, is homeomorphic to G, and so is still trip. conn. One argues readily that it is also trip. conn. in the slightly different sense of this term used by Whitney.[15] Then by Whitney's result[16] it follows that any two vertices of G are joined by three distinct chains, as asserted.

4. **Maps of triply connected graphs.** A *map* σ of a combinatorial graph G is a correspondence which assigns to each vertex p of G a point σp on the sphere and to each edge α of G a Jordan arc $\sigma(\alpha)$ on the sphere, such that the ends of $\sigma(\alpha)$ are the maps of the ends of α, while $\sigma p \neq \sigma q$ if $p \neq q$ and two arcs $\sigma(\alpha)$ and $\sigma(\beta)$ with $\alpha \neq \beta$ do not intersect, except perhaps at their end points. Any subgraph E of G has an image $\sigma(E)$ composed of all those arcs $\sigma(\alpha)$ with α in E. Two maps $\sigma(G)$ and $\tau(G)$ of G on the sphere will be considered as identical if there exists a topological transformation of the sphere carrying $\sigma(G)$ into $\tau(G)$ and each $\sigma(\alpha)$ and $\sigma(p)$ into the corresponding $\tau(\alpha)$ or $\tau(p)$.

If G is non-separable, any map $\sigma(G)$ cuts the sphere into a number of connected domains whose boundaries are the maps of certain circuits of G. These circuits we call the *complementary domain boundaries* (c. d. boundaries) of the map. We state without proof the following

[15] H. Whitney, *Congruent graphs*, loc. cit., p. 158.

[16] H. Whitney, ibid., Theorem 7, p. 160.

THEOREM 4. *Two maps of a cyclically connected graph on the sphere are topologically equivalent if and only if they have the same set of complementary domain boundaries.*[17]

The possible c. d. boundaries of a non-separable graph can be characterized in the following fashion:[18]

THEOREM 5. *A set of circuits $C_1, \cdots, C_m$ in a non-separable graph G is the set of complementary domain boundaries of a planar map of G if and only if each edge of G is contained in exactly two of the circuits C, while the circuits $C_1, \cdots, C_{m-1}$ form a complete independent set*[19] *of circuits in G*, mod 2.

Adkisson has shown[20] that if a cyclically connected graph G has a map in which each pair of c. d. boundaries has a connected or void intersection, then every homeomorphism of G to itself can be extended to the sphere. This suggests that any two maps of such a G on the sphere are identical. But the intersection of any pair of c. d. boundaries will be connected or void if and only if the graph G can not be split. This agrees with a theorem of Whitney,[21] which states that a triply connected graph has at most one dual (and hence at most one map on the sphere). We shall now find an intrinsic characterization of the unique map on the sphere of a triply connected graph.

A circuit C is a *cut circuit* in the triply connected graph G if there are two non-void subgraphs H and H' of G such that $G - C' = H + H'$, while H and H' have in common only vertices on C.

THEOREM 6. *If G is a triply connected graph with a planar map $\sigma(G)$, then a circuit $C \subset G$ is a complementary domain boundary of $\sigma(G)$ if and only if C is not a cut circuit of G.*

Proof. Any C not a c. d. boundary of σ certainly cuts G into the two non-void pieces located respectively within and without the closed curve $\sigma(C)$. Conversely, let C be a c. d. boundary of $\sigma(G)$ and suppose, contrary to the theorem, that C is a cut circuit of G. Then $\sigma G - \sigma C$ is not connected, and so has $m \geqq 2$ connected pieces $E_1, E_2, \cdots, E_m$. Each closure $\bar{E}_i$ consists of E_i plus end points of arcs, and so is the map of a corresponding subgraph G_i of G.

[17] As the number of domain boundaries is finite, this theorem may be readily proved by mathematical induction from the fact that a topological correspondence between any two closed Jordan curves can be extended to the interiors of these curves. The induction could depend on Mac Lane, loc. cit., Theorem 3.1. The theorem also follows by the method used by V. W. Adkisson, loc. cit., Theorem 2; it is based on a theorem due to H. M. Gehman, *On extending a continuous one-to-one correspondence of two plane continuous curves to a correspondence of their planes*, Transactions of the American Mathematical Society, vol. 28 (1926), pp. 252–265.

[18] This is a re-formulation of the condition quoted in the introduction, using Mac Lane, loc. cit., Lemma 4.1, Theorem 5.1, and Theorem 5.3.

[19] The second condition is equivalent to requiring that $C_1, \cdots, C_{m-1}$ are independent modulo 2 and that $m - 1$ is the nullity of G. The nullity of G is $E(G) - V(G) + P(G)$, where $E(G)$, $V(G)$, and $P(G)$ are respectively the number of edges, the number of vertices, and the number of connected pieces of G.

[20] V. W. Adkisson, loc. cit., Theorem 3, p. 168.

[21] H. Whitney, *Congruent graphs*, Theorem 11.

These G_i are the smallest sets into which C cuts G. Those vertices of each G_i which lie on C we call the *feet* of G_i.

Each G_i has at least three feet. For if G_i had no foot on C, G would be disconnected; if G_i had one foot on C, G would be separable at this foot, while if G_i had two distinct feet, G could be split at these two feet into G_i plus the remaining part of G, unless G_i were a single chain. If G_i is a single chain with two feet on C, this chain divides the exterior of C into two regions. The two subgraphs of G contained in the closures of these two regions, respectively, intersect only in the two feet of G_i, so that G is again split. Each of these results contradicts the hypothesis that G is triply connected.

Let p_1 and p_2 be two feet of G_1. Because E_1 is connected, p_1 and p_2 can be joined by a chain L in G_1. Any set $E_i \neq E_1$ contains no points of σL or of the c. d. boundary σC, so that E_i must lie entirely within one of the two regions bounded by σL and an arc of σC. Therefore, the feet of each $G_i \neq G_1$ all lie on one of the two arcs into which p_1 and p_2 divide C.

We can choose the feet p_1 and p_2 of G_1 so that one of the arcs C_1 into which they divide C contains all the feet of G_2, but no feet of G_1 except for p_1 and p_2. Then G can be split. For let H denote the subgraph of G composed of C_1 and all those subgraphs G_i whose feet lie only on C_1, while H' consists of $C - C_1$ and all those G_i whose feet all lie on $C - C_1$. By the previous result, every G_i belongs to exactly one of these subgraphs, so that $G = H + H'$, where H and H' have only p_1 and p_2 in common. As G_1 has at least three feet, G_1 is in H', while G_2 is in H, so that neither H nor H' is a single chain. Therefore this is a split, contrary to the triple connectivity of G. Hence the c. d. boundaries can not cut G, as asserted in the theorem.

THEOREM 7. *A triply connected graph can not have two topologically distinct maps on the sphere.*

Proof. If there is a map, its c. d. boundaries are exactly the circuits which do not cut G. This property is independent of the particular map, so that there can be only one set of c. d. boundaries and hence by Theorem 4 at most one map on the sphere. Another immediate result is the following criterion for mappability:

THEOREM 8. *A triply connected graph G can be mapped on the sphere or on the plane if and only if the circuits of G which do not cut G form a set of circuits satisfying the condition of Theorem 5 for a set of c. d. boundaries.*

5. Maps of cyclically connected graphs. The final theorem on mappability is

THEOREM 9. *A cyclically connected graph G can be mapped on the plane if and only if all its atoms can be mapped on the plane; that is, if and only if all its triply connected atoms satisfy the mappability criterion of Theorem 8.*

This theorem will follow by induction on the number of atoms once we show that a graph G split into two blocks has a map whenever both blocks have maps. This fact we state in the following more explicit form.

LEMMA 3. *If G is split as in* (2), *and if τ and τ' are maps on the sphere of the*

blocks $H + X$ and $H' + X'$ respectively, where in τ the chain X appears in two c. d. boundaries C and D, while X' appears in τ' in the boundaries C' and D', then there is a map of G in which the c. d. boundaries are the c. d. boundaries of τ and τ', except that the circuits C, D, C', and D' are replaced by

$$(6) \qquad (C - X) + (C' - X'), \qquad (D - X) + (D' - X').$$

Proof. Because the chains $C - X$ and $C' - X'$ belong to H and H' respectively, they have no arcs in common, while by construction they both have the two vertices h_1 and h_2 as ends. Hence $(C - X) + (C' - X')$ and likewise the other graph in (6) is actually a circuit in G. Once the lemma is established, a simple change of notation will give a similar result with (6) replaced by

$$(7) \qquad (C - X) + (D' - X'), \qquad (D - X) + (C' - X').$$

To prove the lemma we need only squeeze the map $\tau(H')$ into the region of the map $\tau(H)$ originally occupied by $\tau(X)$. This may be done as follows.

Draw an arc $\tau(Z)$ with ends $\tau(h_1)$ and $\tau(h_2)$ in the region bounded by C and draw a similar arc $\tau(Z')$ inside C'. In these altered maps the c. d. boundaries are as before except that C and C' are replaced by

$$(8) \qquad Z + (C - X), \quad Z + X, \qquad Z' + (C' - X'), \quad Z' + X',$$

respectively. One of the two regions of the sphere bounded by $X + Z$ contains the rest of $\tau(H)$. Call this region the "outside" of $X + Z$, with a similar convention as to $X' + Z'$. Then map the outside and boundary of $X' + Z'$ topologically on the inside and boundary of $X + Z$ in such a fashion that $\tau'(h_1)$ and $\tau'(h_2)$ go into $\tau(h_1)$ and $\tau(h_2)$ respectively, while $\tau'(X')$ goes into $\tau(X)$ and $\tau'(Z')$ into $\tau(Z)$. The two maps so combined have c. d. boundaries which are the original c. d. boundaries of τ and τ' except that C, C', D, and D' are replaced by

$$(9) \quad Z + (C - X), \quad Z + (C' - X'), \quad X + (D - X), \quad X + (D' - X').$$

If we remove the maps of X and Z, a map of G with the required c. d. boundaries will remain. The lemma being established, the theorem follows.

6. **The number of maps of a graph.** In this section we denote by $\mu(G)$ the number of topologically distinct maps of a graph G on a sphere. If one map of G is given, other maps may be found by "rotating" one of the components of a split of G or by permuting several components which are connected in "parallel". To show that all distinct maps of G can be obtained by sequences of such operations, we first discuss such components in "parallel". Given two vertices p and q of G, we say that two edges α and β of G are *connected outside* of p and q if there is a chain L of G containing α and β but passing through neither p nor q. The relation, "α is connected to β outside p and q", is reflexive, symmetric, and transitive. Therefore G is subdivided into disjoint subgraphs $M_1, M_2, \cdots, M_t$

such that two edges of G belong to the same subgraph if and only if they are connected outside p and q. Consequently,

$$G = M_1 + M_2 + \cdots + M_t \tag{10}$$

is a representation of G in which any two of the subgraphs M_i and M_j have only the vertices p and q in common. Furthermore, no subgraph M_i has a semi-split at p and q, so that the representation (10) can not be further subdivided. If $t > 2$ in (10), we say that G has a *multiple split* of *order* t at p and q.

THEOREM 10. *If G can be mapped on the sphere, the number $\mu(G)$ of topologically distinct maps of G on the sphere is*

$$\mu(G) = 2^{\alpha-1} \prod_{i=1}^{k} (t_i - 1)!, \tag{11}$$

where $\alpha = \alpha(G)$ is the number of atoms in a complete set of atoms of G, where G has k distinct multiple splits, and where the i-th multiple split has the order t_i.

We shall establish this theorem by constructing the atoms from "least" splits of G. A component H in a split (1) of G at h_1 and h_2 will be called *least* if the component H has no semi-split at h_1 and h_2; that is, if any two arcs of H are connected outside h_1 and h_2. A split (1) of G will itself be called *least* if either of the components H or H' of the split is least.

LEMMA 4. *Let* (1) *be a least split of G into H and H'. Then in any map σ of G there are two c. d. boundaries which have edges in both H and H', while any other c. d. boundary is entirely in H or entirely in H'.*

Proof. Edges of both H and H' abut on the split point h_1. The cyclic order of the complementary domains about h_1 indicates that through this point there must pass at least two c. d. boundaries which have edges in both H and H'.

Suppose there were three (or more) such boundaries with edges in both H and H'. If the three complementary domains with these boundaries are removed, the remaining part of the sphere falls into three parts which touch only at h_1 and h_2. Since H has edges in all three c. d. boundaries, it follows readily that H must have edges in at least two of the three parts of the sphere. As these two parts of H touch only in h_1 and h_2, H has a semi-split at these vertices. H' likewise has a semi-split, contrary to the assumption that one of H and H' is least. Therefore there are but two c. d. boundaries of the sort considered.

LEMMA 5. *If G has a least split* (2), *there exists a* 2-1 *correspondence between the maps σ of G on the sphere and the pairs of maps (τ, τ'), where τ and τ' are maps on the sphere of the blocks $H + X$ and $H' + X'$ respectively.*

Proof. A given map σ of the whole graph G is also a map of each of the subgraphs (or blocks) $H + X$ and $H' + X'$, and hence does correspond to a pair of maps of these blocks. Specifically, let σ be determined (Theorem 4) by its c. d. boundaries

$$C_1, C_2, \cdots, C_k, \qquad C'_1, \cdots, C'_m, \qquad L + L', \qquad M + M', \tag{12}$$

where each C_i is a circuit contained in H, and each C'_j is contained in H', while,

as in the last lemma, the two circuits with edges in both H and H' consist of chains L and M in H and L' and M' in H'. The sub-map of $H + X$ arises by dropping from the map (12) all edges of $H' - X$. The first circuits C_1 to C_k must thus remain as c. d. boundaries in the map of $H + X$, while by Theorem 5 there must be two other c. d. boundaries containing all of L and all of M respectively. Furthermore, X must be in two boundaries, so that the additional boundaries are just $X + L$ and $X + M$. This gives a map of $H + X$ with the c. d. boundaries

$$C_1, C_2, \cdots, C_k, \qquad L + X, \qquad M + X, \tag{13}$$

and similarly a map of the other block $H' + X'$ with the boundaries[22]

$$C_1', C_2', \cdots, C_m', \qquad L' + X', \qquad M' + X'. \tag{14}$$

Therewith we have a correspondence between the map (12) and the "pair of maps" (13) and (14). Conversely, given any two maps (13) and (14) of the two blocks, there can be at most two maps of G corresponding to (13) and (14) in this fashion; namely, the maps given by the c. d. boundaries in (12) or in

$$C_1, C_2, \cdots, C_k, \qquad C_1', \cdots, C_m', \qquad L + M', \qquad M + L'. \tag{15}$$

Both of these maps (12) and (15) are geometrically possible by Lemma 3. Hence the correspondence is a 2–1 correspondence, as asserted in Lemma 5.

LEMMA 6. *If G has a least split* (1), *then, for any multiple split of G of order t at p and q, one of the blocks* (2) *of the least split has a multiple split of order t at p and q, while the other block has no multiple split at p and q. The splits so obtained are the only multiple splits of the blocks.*

The second statement is simple, for if one of the blocks has a multiple split at the vertices r and s, then, by the definition of a split in (10), the original graph G also has a multiple split at r and s. The proof of the first part depends essentially on the fact that one of the components, say H, of the given split (1) is least.

Case 1. $p = h_1$, and $q = h_2$. As H is least, it must be one of the M_i in the given multiple split (10) of G, say M_1. Then

$$H' = M_2 + \cdots + M_t,$$

so that the other block $H' + X'$ has a multiple split like (10) with M_1 replaced by X', and this multiple split has the order t.

Case 2. Neither H nor H' contains both p and q, so that one vertex, say p, is in H but not H', while q is in H', but not in H. But in the multiple split (10) each M_i must, because G is cyclically connected, contain a chain with ends p and q, and this chain must pass through one of the split vertices h_1 or h_2 of the split (1). Since each of these split vertices is in but one of the M's, there can

[22] This could also be proved purely combinatorially by showing that if the set of circuits (12) satisfies Theorem 5, the two sets (13) and (14) also satisfy Theorem 5.

thus be only two subgraphs M_i, and the presumed decomposition (10) is no honest multiple split.

There remains the case when p and q are both in one component, say the component H of (1), while one of the split vertices, say h_1, is at neither p nor q. Then any arc of H' is connected to this h_1 by a chain not passing through either p or q, so that all of H' is contained in one of the pieces M_i of the given multiple split (10) of G. Thus, when all these edges of H' are replaced simply by X', (10) gives a multiple split of the block $H + X$ with the same order t, while the other block $H' + X'$ can not contain both p and q as branch vertices and so certainly has no multiple split at p and q.

In all cases, the given multiple split yields a similar multiple split of just one of the blocks (2), so that the lemma holds.

Lemma 7. *If a graph G has no least split, then G is triply connected or is a branch graph or a circuit.*

For if G is not triply connected, it must have a split at some pair of vertices (p, q), and hence a representation (10). There must be $t \geqq 3$ terms, because the split is not least. If one of the subgraphs M_i were not a single arc, then $G = M_i + (G - M_i)$ would be a split in which M_i is least. Such a least split is impossible by hypothesis, so each M_i is a single arc, and G is a branch graph.

To prove Theorem 10, decompose G by successive least splits until none of the resulting blocks have least splits. By Lemma 7 these blocks are either branch graphs or else are triply connected and hence atoms of G, so that we have a set of blocks

$$A_1, A_2, \cdots, A_\alpha, \qquad B_1, B_2, \cdots, B_\beta, \tag{16}$$

where the α graphs A_i form a complete set of atoms, while the B_i are branch graphs. By Lemma 6 each multiple split of G corresponds to one and just one multiple split with the same order in one of the blocks (16). Of these blocks, only the branch graphs B_i have multiple splits, while G has multiple splits of order t_i, $i = 1, \cdots, k$. Hence the branch graphs in (16) can be so arranged that B_i is a branch graph with t_i branches. There are $\beta = k$ such graphs.

In the set of subgraphs (16) each triply connected atom A_i has but one map on the sphere, while each branch graph B_i with t_i branches has $(t_i - 1)!/2$ different maps (this is the number of ways of arranging the t_i branches in cyclic order). We combine the maps of (16) to get maps of G. Each step in the combination will by Lemma 5 yield two alternative maps of G, while there are $\alpha + k - 1$ combinations in all. Hence there are

$$2^{\alpha+k-1} \prod_{i=1}^{k} \frac{(t_i - 1)!}{2} = 2^{\alpha-1} \prod_{i=1}^{k} (t_i - 1)!$$

different maps of G, as asserted in the theorem.

Any planar map of G can be obtained by stereographically projecting a map from the sphere onto the plane. Distinct maps are obtained when the north

pole is chosen in distinct regions, and the number of such regions is simply $N(G) + 1$; therefore we have the

COROLLARY. *The number of topologically distinct maps of a cyclically connected planar graph on the plane is*

$$\nu(G) = [N(G) + 1]\mu(G) = [N(G) + 1]2^{a-1} \prod_{i=1}^{k} (t_i - 1)!,$$

where $N(G)$ is the nullity of G and the other constants are given as in the theorem.

CORNELL UNIVERSITY.

CRITICAL CURVATURES IN RIEMANNIAN SPACES

By Arthur B. Brown

1. **Introduction.** A well known theorem in differential geometry concerns the normal curvatures of curves through a point on a 2-dimensional surface in 3-space. It states that either the curvature is constant, independent of the direction of the curve, or else there is one direction giving a maximum to the curvature, and another (perpendicular) direction giving a minimum. We generalize this result to the case of an n-surface in a Riemannian $(n + 1)$-space. In place of merely a maximum and a minimum, there is in general a non-degenerate critical point of each type or index[1] $0, 1, \cdots, n - 1$. A similar result is obtained for an arbitrary subspace of a Riemannian space, the theorem being stated in terms of projections on any direction orthogonal to the subspace. A final theorem, with a similar statement regarding critical values, concerns the Ricci mean curvature in a Riemannian space.

2. **The principal directions[2] for a real quadratic form.** Our results regarding critical values will be based on the following theorem.

Theorem 2.1. *Given the real quadratic form*[3]

$$z = a_{ij}x_ix_j, \tag{2.1}$$

on the locus

$$x_ix_i = 1 \tag{2.2}$$

z has at most, and in general exactly, n distinct critical values. When the number is n, the critical values are taken on at n pairs of diametrically opposite points of (2.2), *determining n mutually perpendicular lines through the origin in the number space of the x's. If the pairs are ordered according to the algebraic values of z, at either point of the i-th pair z has a non-degenerate critical point of index $i - 1$.*

Proof. We begin by making an orthogonal transformation with fixed origin in (x)-space so that the given form becomes (using the same letters $x_1, \cdots, x_n$)

$$z = b_1x_1^2 + \cdots + b_nx_n^2 = b_ix_i^2 \tag{2.3}$$

with the b's real.[4] Now if we consider the function

Received February 24, 1937; presented to the American Mathematical Society, March 26, 1937.

[1] Marston Morse, *The Calculus of Variations in the Large*, Amer. Math. Soc. Colloquium Publications, vol. 18, p. 143.

[2] Cf. L. P. Eisenhart, *Riemannian Geometry*, Princeton, 1926, p. 110. We shall refer to this volume as Eisenhart.

[3] Repetition of an index indicates summation from 1 to n.

[4] Cf. M. Bôcher, *Introduction to Higher Algebra*, p. 170, Theorem 1 and p. 171, Theorem 2; or L. E. Dickson, *Modern Algebraic Theories*, p. 74, Theorem 10.

$$w = b_i x_i^2 / x_j x_j, \tag{2.4}$$

we see that $w = z$ on locus (2.2). Using the fact that w is constant along each line directed towards the origin, omitting the origin itself, we easily see that those of the critical points of w, as a function of n independent variables, which are located on (2.2), are the same points as the critical points of z on (2.2) as a function of $n - 1$ independent variables.

Then to find the critical points, using (2.4) we set

$$0 = \frac{\partial w}{\partial x_s} = \frac{(x_j x_j)(2b_s x_s) - (b_i x_i^2)(2x_s)}{(x_j x_j)^2} \qquad \text{(do not sum } s\text{)}.$$

Hence the critical points of z on (2.2) are the points on (2.2) where

$$b_s x_s = (b_i x_i^2) x_s \qquad (s = 1, \cdots, n; \text{ do not sum } s). \tag{2.5}$$

Obviously

$$x_j = \delta_{kj} \qquad (j = 1, \cdots, n)^{5} \tag{2.6}$$

satisfy (2.2) and (2.5) for any fixed k. Taking $k = 1, 2, \cdots, n$, we have n solutions of (2.5) and (2.2).

We shall now show that (α) the number of distinct critical values of z on (2.2) equals the number of distinct b's in (2.3), and (β) if that number is n, then (2.6) gives the only critical points other than those with $x_j = -\delta_{kj}$.

Suppose

$$b_1 = b_2 = \cdots = b_r, \tag{2.7}$$

but no other $b = b_1$. If we take a solution of (2.2) and (2.5) with one or more of $x_1, x_2, \cdots, x_r$ different from zero, say $x_1 \neq 0$, then if, say, $x_{r+1} \neq 0$, from (2.5) with $s = 1$ and $s = r + 1$ we would have

$$b_i x_i^2 = b_1 = b_{r+1},$$

contrary to hypothesis. Hence, for the solution in question,

$$0 = x_{r+1} = \cdots = x_n \tag{2.8}$$

and hence

$$1 = x_1^2 + \cdots + x_r^2. \tag{2.9}$$

From (2.7), (2.8), (2.9) and (2.3) we now see that the critical value in question is $b_1 = b_2 = \cdots = b_r$. Since a similar argument holds for each set of equal b's, our assertion (α) follows at once.

If we have n distinct critical values, then by (α) the number of distinct b's must be n, and hence for each critical point only one of the x's can be different from zero (cf. (2.8)). Thus (β) is established. The perpendicularity of the directions also follows.

[5] $\delta_{kj} = 0$ if $k \neq j$, $= 1$ if $k = j$.

Consider the solution (2.6) with $k = 1$. It is $(1, 0, \cdots, 0)$. At this point $x_2, \cdots, x_n$ can be taken as the independent variables for (2.2).[6] Substituting for x_1^2 from (2.2) into (2.3), we have

$$z = b_1(1 - x_2^2 - \cdots - x_n^2) + b_2 x_2^2 + \cdots + b_n x_n^2 \tag{2.10}$$
$$= b_1 + (b_2 - b_1)x_2^2 + (b_3 - b_1)x_3^2 + \cdots + (b_n - b_1)x_n^2 .$$

Hence the index of the critical point (the number of negative coefficients) is the number of b's less than b_1.

Since a similar result can be obtained by using (2.6) with each of the values $k = 2, 3, \cdots, n$, we infer that if, say, $b_1 < b_2 < \cdots < b_n$, then, for each k, $x_j = \delta_{kj}$ gives us a critical point of index $k - 1$, at which $z = b_k$. We would obtain the same result by taking $x_j = -\delta_{kj}$. The truth of the theorem is now established.

3. **Riemannian coördinates.** In this section we establish Riemannian coördinates[7] for a Riemannian space without assuming that the coefficients of the fundamental form are analytic. If one is satisfied with the case that the coefficients are analytic, this section may be omitted.

THEOREM 3.1. *If the fundamental form of a Riemannian space has coefficients of*[8] *class C^k, $k \geqq 4$, neighboring a point P with coördinates (a), a Riemannian coördinate system can be introduced, for a neighborhood of P, with origin at P, by a non-singular transformation of coördinates in terms of functions of class C^{k-1}.*[9]

Proof. If we take $x_1, \cdots, x_n$ as the coördinates, and $g_{ij}dx_i dx_j$ as the fundamental form, the geodesics are the solutions of

[6] The index of the critical point is independent of the particular parameters chosen, providing z is a function of class C^2 of those parameters.

[7] A system of coördinates $y_1, \cdots, y_n$ in a Riemannian n-space is called *Riemannian* if the geodesics through the origin are the curves given by the equations $y_i = \lambda^i t$ $(i = 1, \cdots, n; \lambda^1, \cdots, \lambda^n$ any real constants not all zero). The theorem of this section has been proved by J. H. C. Whitehead, *On the covering of a complete space by geodesics through a point*, Annals of Math., vol. 36 (1935), pp. 679–704. Cf. footnote 5 of T. Y. Thomas, *On normal coördinates*, Proc. Nat. Acad. Sci., vol. 22 (1936), pp. 309–312, where a proof of the theorem is based on results in an earlier paper by W. Mayer and T. Y. Thomas, Math. Zeitschrift, vol. 40 (1936), pp. 658–661. The proof was also given in some mimeographed notes of W. Mayer at Princeton in 1936, for a more general variational problem. We are indebted to G. Comenetz for the above information. Since a simple proof has not yet been published all in the same paper, we think it advantageous to give one here. In the papers cited above, the hypotheses are equivalent to taking $k \geqq 2$ in Theorem 3.1 instead of $k \geqq 4$. We assume $k \geqq 4$ to insure that the equations of the geodesics in the new coördinate system have coefficients of class C^1.

[8] A function is of class C^k if it is continuous and has all its partial derivatives, up to and including those of order k, continuous.

[9] While we shall be dealing with the case that the fundamental form is positive definite, we do not make this assumption in this theorem.

$$\frac{d^2x_i}{dt^2} + \begin{Bmatrix} jk \\ i \end{Bmatrix} \frac{dx_j}{dt} \frac{dx_k}{dt} = 0 \qquad (i = 1, \cdots, n).\tag{3.1}$$[10]

Since $\begin{Bmatrix} jk \\ i \end{Bmatrix}$ is of class C^{k-1}, the solution with $x_i = x_i^0$ and $(dx_i/dt) = (dx_i/dt)_0$ at $t = t_0$ is given, according to a theorem of differential equations, by

$$\begin{aligned} (3.2) \quad x_i &= \phi_i[t, t_0, x_1^0, \cdots, x_n^0, (dx_1/dt)_0, \cdots, (dx_n/dt)_0] \\ &\equiv \phi_i[t, t_0, x^0, (dx/dt)_0] \qquad (i = 1, \cdots, n), \end{aligned}$$

with ϕ_i of class C^{k-1} in all the arguments,[11] say for $|t| < \epsilon$, $|t_0| < \epsilon$, $|x_i^0 - a_i| < \epsilon$, $|(dx_i/dt)_0| < \epsilon$, $\epsilon > 0$. Taking $x_i^0 = a_i$ and $t_0 = 0$, we have the geodesics through (a) represented by

$$x_i = \phi_i[t, 0, a, (dx/dt)_0] \qquad (i = 1, \cdots, n),\tag{3.3}$$

with ϕ_i of class C^{k-1} for $|t| < \epsilon$ and $|(dx_i/dt)_0| < \epsilon$. Note that, since $x_i = a_i$ is a solution of (3.1), from the uniqueness of the solution we have

$$\phi_i(t, 0, a, 0) \equiv a_i \qquad (i = 1, \cdots, n), \ |t| < \epsilon.\tag{3.4}$$

Now let us make a change of independent variable, $\bar{t} = ct$, $c \neq 0$, constant, replacing (3.1) by

$$\frac{d^2x_i}{d\bar{t}^2} + \begin{Bmatrix} jk \\ i \end{Bmatrix} \frac{dx_j}{d\bar{t}} \cdot \frac{dx_k}{d\bar{t}} = 0 \qquad (i = 1, \cdots, n).\tag{3.5}$$

Since (3.5) is of the *same form as* (3.1), the solutions of (3.5) through (a) and with $\bar{t}_0 = 0$ are given by

$$\begin{aligned} (3.6) \quad x_i &= \phi_i[\bar{t}, 0, a, (dx/d\bar{t})_0] \\ &= \phi_i[ct, 0, a, (dx/dt)_0/c], \qquad |ct| < \epsilon, \qquad |(dx/dt)_0/c| < \epsilon. \end{aligned}$$

Since (3.5) is *equivalent to* (3.1), we have, from (3.3) and (3.6),

$$\phi_i[t, 0, a, (dx/dt)_0] \equiv \phi_i[ct, 0, a, (dx/dt)_0/c] \qquad (i = 1, \cdots, n),\tag{3.7}$$

if every argument except the a's is less than ϵ in absolute value.[12]

[10] $\begin{Bmatrix} jk \\ i \end{Bmatrix} = g^{ih}\begin{bmatrix} jk \\ h \end{bmatrix} = \frac{g^{ih}}{2}\left(\frac{\partial g_{jh}}{\partial x_k} + \frac{\partial g_{kh}}{\partial x_j} - \frac{\partial g_{jk}}{\partial x_h}\right)$, if we use the customary notation. Cf. Eisenhart, p. 50, equation (17.8) and the proof following that equation. In the proof, in place of covariant derivatives with respect to x^k multiplied by dx^k/ds and summed for k, absolute derivatives with respect to s should be used, as the functions differentiated are not functions of the x's. Cf. our footnote 17.

[11] See G. A. Bliss, *Solutions of differential equations of the first order as functions of their initial values*, Annals of Math., (2), vol. 6 (1905), pp. 49–68 (theorem on p. 67); or Bolza, *Variationsrechnung*, 1909, p. 178.

[12] The work is carried this far in Duschek-Mayer, *Lehrbuch der Differentialgeometrie*, vol. 2, pp. 93–94, but it is not completed there. Cf. footnote 2 on page 95 there. We shall refer to this volume as Mayer. (Vol. 2 is by W. Mayer.) See also G. A. Bliss and Max Mason, *Fields of extremals in space*, Trans. Amer. Math. Soc., vol. 11 (1910), pp. 325–340.

Now place the restriction $|t| < \eta = \epsilon/2$. For any such $t \neq 0$, if $c = \epsilon/2t$, all the arguments in (3.7) except the a's will be numerically less than ϵ provided $|(dx_i/dt)_0| < \epsilon$. Hence

$$\begin{aligned}(3.8)\qquad \phi_i[t, 0, a, (dx/dt)_0] &\equiv \phi_i[\epsilon/2, 0, a, 2t(dx/dt)_0/\epsilon] \\ &\equiv \psi_i[t(dx_1/dt)_0, \cdots, t(dx_n/dt)_0]\end{aligned}$$

for $0 < |t| < \eta$, $|(dx_i/dt)_0| < \eta$. When $t = 0$ the left-hand member of (3.8) is a_i. The second member is also a_i, by (3.4), when $t = 0$. Hence (3.8) holds also when $t = 0$, and thus the solutions of (3.1) are given by

$$(3.9)\qquad x_i = \psi_i[t(dx_1/dt)_0, \cdots, t(dx_n/dt)_0] \qquad (i = 1, \cdots, n)$$

with ψ_i of class C^{k-1} for $|t| < \eta$, $|(dx_i/dt)_0| < \eta$.

Now differentiate each side of (3.9) with respect to t and set $t = 0$. Denoting by $y_1, \cdots, y_n$ the arguments of ψ_i, we have

$$(3.10)\qquad (dx_i/dt)_0 \equiv [\partial\psi_i(0)/\partial y_j](dx_j/dt)_0 \qquad (i = 1, \cdots, n).$$

Since (3.10) are identities in the $(dx_i/dt)_0$'s, we infer

$$(3.11)\qquad \partial\psi_i(0)/\partial y_j = \delta_{ij} \qquad (i, j = 1, \cdots, n).$$

Now consider the transformation

$$(3.12)\qquad x_i = \psi_i(y_1, \cdots, y_n) \qquad (i = 1, \cdots, n).$$

From (3.11) we see that, for some $\zeta < \eta\cdot\eta = \eta^2$, $\zeta > 0$, this transformation is one-to-one for $|y_i| < \zeta$. Since the ψ_i are of class C^{k-1}, the new fundamental form $g'_{ij}dy_i dy_j$ will have coefficients of class C^{k-2}; and since $k - 2 \geqq 2$, the geodesics are uniquely determined in the new coördinate system.

Now consider the equations

$$(3.13)\qquad y_i = \lambda^i t \qquad (i = 1, \cdots, n),$$

where $\lambda^1, \cdots, \lambda^n$ are (real) constants with $|\lambda^i| < \eta$ and $0 < \lambda^i\lambda^i$. The curve (3.13) is given in the (x)-system by

$$(3.14)\qquad x_i = \psi_i[\lambda^1 t, \cdots, \lambda^n t], \qquad |t| < \eta.$$

Hence, comparing with (3.8), we see that it is the geodesic through (a) in direction $(\lambda^1, \cdots, \lambda^n)$. We now see that the restriction $|\lambda^i| < \eta$ can be dropped, provided that for each set (λ) we restrict t sufficiently. Thus all the geodesics through P are the curves given in the new coördinate system by equations of the form (3.13) with $0 < \lambda^i\lambda^i$. It follows that the new system is Riemannian, with origin at P, and the proof is complete.

4. **Hypersurfaces in a Riemannian space.** Here we take the case of a surface of dimension one less than that of the space.

Definition. The locus of a set of simultaneous equations

$$\phi_i(x) \equiv \phi_i(x_1, \cdots, x_m) = 0 \qquad (i = 1, \cdots, k; k < m)$$

is called a *regular* $(m - k)$*-spread of class* C^r neighboring a point $(x_1^0, \cdots, x_m^0)$ in a Euclidean or Riemannian space with coördinates (x), if (x^0) lies on the locus, the functions ϕ_i are of class C^r near (x^0), $r \geqq 1$, and the matrix of the first partial derivatives is of rank k at (x^0).

Theorem 4.1. *If P is a point on a regular n-spread S, $n > 1$, of class C^2 in a Riemannian $(n + 1)$-space R with positive definite fundamental form having coefficients of class C^4, the normal curvatures*[13] *of curves on S through P, as functions of parameters determining their directions, have at most n critical values. When, as is in general the case, the number is n, the values are taken on in n mutually perpendicular directions;*[14] *if the critical values are ordered according to their algebraic values, then the i-th is at a critical point of index $i - 1$.*

Proof. Using Theorem 3.1, we introduce Riemannian normal coördinates[15] with origin at P in such a way that the surface $x_{n+1} = 0$ is tangent at P to S. Then, if we denote $x_{n+1} = z$, S is given by

$$z = f(x_1, \cdots, x_n), \qquad f \text{ of class } C^2, \tag{4.1}$$

with

$$0 = \frac{\partial f}{\partial x_1}(0, \cdots, 0) = \cdots = \frac{\partial f}{\partial x_n}(0, \cdots, 0). \tag{4.2}$$

Now let any unit vector (λ) be given at P, so that

$$\lambda^i \lambda^i = 1. \tag{4.3}$$

Let

$$\begin{aligned} x_i &= \phi_i(s) \qquad (i = 1, \cdots, n), \\ z &= f[\phi_1(s), \cdots, \phi_n(s)] \end{aligned} \tag{4.4}$$

be a curve, with s the arc length, $s = 0$ at P, such that

[13] For a curve on S through P with non-zero first curvature, the vector in the direction of the principal normal and of length equal to the absolute value of the curvature will be called the *curvature vector*. If the curve has first curvature zero at P, the curvature vector is the vector with all components zero. The *normal curvature* of a curve through P on S is the projection on the normal direction to S at P, with a sense assigned to the latter, of the curvature vector. All regular curves of class C^2 on S tangent to a given curve on S at P have the same curvature vector. Cf. Eisenhart, p. 151. See also our Theorem 5.1 and the definition preceding it. For a curve whose curvature vector is orthogonal to S, the normal curvature is plus or minus the absolute value of the actual first curvature.

[14] This is well known. Cf. Eisenhart, p. 153.

[15] Riemannian normal coördinates are Riemannian coördinates such that $g_{ij} = \delta_{ij}$ at the origin (Eisenhart, p. 55). They are easily obtained from any Riemannian coördinates, when the fundamental form is positive definite, by a linear transformation. The notation in Mayer is slightly different.

$$\phi_i'(0) = \lambda^i \qquad (i = 1, \cdots, n).\text{[16]} \tag{4.5}$$

Since we have normal coördinates, for curve (4.4) at P we have

$$\frac{\delta}{\delta s}\frac{dx_\alpha}{ds} = \frac{d}{ds}\frac{dx_\alpha}{ds} = \frac{d^2 x_\alpha}{ds^2} \qquad (\alpha = 1, \cdots, n + 1).\text{[17]}$$

Hence the curvature vector of the curve at P has its α-th component[18] d^2x_α/ds^2 $(\alpha = 1, \cdots, n + 1)$. Hence the normal curvature, projection of the curvature vector on the direction of the positive z-axis, is $d^2x_{n+1}/ds^2 = d^2z/ds^2$. Now

$$\frac{dz}{ds} = \frac{\partial f}{\partial x_i}\cdot\frac{d\phi_i}{ds},$$

and therefore at P the normal curvature

$$K = \frac{d^2 z}{ds^2} = \frac{\partial^2 f(0)}{\partial x_i \partial x_j}\frac{d\phi_i}{ds}\cdot\frac{d\phi_j}{ds} = \frac{\partial^2 f(0)}{\partial x_i \partial x_j}\lambda^i\lambda^j, \tag{4.6}$$

because of (4.2). Hence

$$K = a_{ij}\lambda^i\lambda^j, \qquad a_{ij} = \frac{\partial^2 f(0)}{\partial x_i \partial x_j}. \tag{4.7}$$

Theorem 4.1 now follows from Theorem 2.1.

We observe that, as a special case, the $(n + 1)$-space may be Euclidean.

5. **Subspaces of a Riemannian space.** In this section we take the case of an n-dimensional surface S in a Riemannian m-space, $m > n$, dropping the restriction that $m = n + 1$. The idea, used here, of projecting the curvature vector on a direction orthogonal to S is found in Mayer,[19] used in another connection.

Let a Riemannian m-space R be given, with positive definite fundamental form having coefficients of class C^4, the variables being $y_1, \cdots, y_m$. Let S be a regular n-spread of class C^2 neighboring a point P of R. Then S is itself a Riemannian n-space, whose fundamental form has coefficients of class C^1.[20]

[16] For example, we could take the curve $x_i = \lambda^i t$ $(i = 1, \cdots, n)$, $z = f[\lambda^1 t, \cdots, \lambda^n t]$. Then at P the quantities dx_α/dt $(\alpha = 1, \cdots, n + 1)$ are $(\lambda^1, \cdots, \lambda^n, 0)$. If now we change the parameter to arc length s, measured in the direction in which t increases, the quantities dx_α/ds are proportional to dx_α/dt, and also have the sum of their squares equal to 1 at P; hence they are also $(\lambda^1, \cdots, \lambda^n, 0)$. Hence $dx_i/ds = \lambda^i$ $(i = 1, \cdots, n)$, as was to be proved.

[17] The letter δ indicates absolute (covariant) differentiation. Cf. Mayer, p. 31 ff. At the origin in Riemannian coördinates, absolute first derivatives equal the usual first derivatives. Cf. Eisenhart, p. 56, for derivatives with respect to the space coördinates. For derivatives to any parameter (e.g., as s above), the result follows immediately from the fact that at the origin the first partial derivatives of the coefficients of the fundamental form are all zero (Eisenhart, p. 55; Mayer, p. 117).

[18] Cf. Eisenhart, p. 61; Mayer, pp. 59–62.

[19] Pp. 159–160.

[20] If we want S to be sufficiently regular so that the geodesics on S exist and are unique, we can demand that S be of class C^3. Its fundamental form will then have coefficients of class C^2.

We begin by establishing a property of curvature, following the next definition.

DEFINITION. The projection of the curvature vector of a curve in a direction normal to the curve at a point on it will be called the *normal curvature* of the curve for the given direction.

THEOREM 5.1. *All regular curves of class C^2 on S tangent to a given curve of S at a fixed point P have the same normal curvature for any direction normal to S.*

Proof. Introduce Riemannian normal coördinates in R so that the locus $0 = y_{n+1} = \cdots = y_m$ becomes tangent to S at the origin P. As the new coördinates are obtained by a transformation using functions of class C^2, the locus S is now given by

$$(5.1) \qquad y_k = f_k(y_1, \cdots, y_n), \qquad f_k \text{ of class } C^2, \qquad (k = n+1, \cdots, m),$$

where all the first partial derivatives are zero at $(0, \cdots, 0)$. Let

$$(5.2) \qquad \begin{aligned} y_i &= p_i(s) && (i = 1, \cdots, n), \\ y_k &= f_k[p_1(s), \cdots, p_n(s)] && (k = n+1, \cdots, m) \end{aligned}$$

be one of the curves in question, with s the arc length. It is easily verified that the p's must be of class C^2.

Since we have normal coördinates, as in §4 the curvature vector has α-th component d^2y_α/ds^2 $(\alpha = 1, \cdots, m)$. Hence its projection on the direction $(0, \cdots, 0, 1)$ is d^2y_m/ds^2. Now

$$\frac{dy_m}{ds} = \sum_{i=1}^{n} \frac{\partial f_m}{\partial y_i} \frac{dp_i}{ds}$$

and therefore, at P,

$$(5.3) \qquad \frac{d^2y_m}{ds^2} = \sum_{i,j=1}^{n} \frac{\partial^2 f_m}{\partial y_i \partial y_j} \frac{dp_i}{ds} \frac{dp_j}{ds},$$

since $\partial f_m/\partial y_i = 0$ at P. As this answer depends only on the direction of the curve (5.2), and since any direction normal to S can be made the direction $(0, \cdots, 0, 1)$, we infer the validity of the theorem.

We now return to the principal question.

THEOREM 5.2. *Let P be a point on a regular n-spread S of class C^2 in a Riemannian m-space, $1 < n < m$, with positive definite fundamental form having coefficients of class C^1. If normal curvatures are taken at P for any given normal direction to S at P, the conclusion of Theorem* 4.1 *holds.*

Proof. Choose coördinates as above, making the given direction the direction $(0, \cdots, 0, 1)$. Since (5.3) is the same kind of result as (4.6), the brief proof is the same as that following (4.6). Hence the theorem is true.

Remark 1. All regular curves on S tangent to a given curve on S and with curvature vectors orthogonal to S or having all components zero have the same curvature vector.

This is a known property,[21] which we state for convenience in reference. It also follows easily from Theorem 5.1.[22]

Remark 2. If, under the hypotheses of Theorem 5.2, S being supposed of class C^3, all the geodesics of S through P have their curvature vectors confined to two opposite directions when non-null, then, in Theorem 5.2, the normal curvatures, for either of those directions, of all curves whose curvature vectors are orthogonal to S are plus or minus the absolute values of the curvatures.

This follows from Remark 1 and the fact that when a geodesic has a non-null curvature vector, the latter is orthogonal to S.[23]

6. **The Ricci mean curvature.** The sum of the $n - 1$ Riemannian curvatures determined by a direction $(\lambda^1, \cdots, \lambda^n)$ in a Riemannian n-space R and each of $n - 1$ other vectors such that all n are mutually orthogonal is called the mean curvature, ρ, of the space for the given direction, and is given by

$$\rho = -\frac{R_{ij}\lambda^i\lambda^j}{g_{ij}\lambda^i\lambda^j}, \tag{6.1}$$

where R_{ij} and g_{ij} are components of the Ricci tensor and the fundamental tensor respectively.[24] If, as in §4, we introduce Riemannian normal coördinates with origin at a given point P where we are considering (6.1), and take (λ) as a unit vector, at P (6.1) reduces to

$$\rho = -R_{ij}\lambda^i\lambda^j. \tag{6.2}$$

We can now apply Theorem 2.1. This gives us the following

THEOREM 6.1. *If P is a point in a Riemannian n-space with positive definite fundamental form having coefficients of class C^4, the mean curvatures at P, as functions of parameters determining the direction, satisfy the conclusion of Theorem* 4.1.

Appendix—The critical diameters of central quadrics[25]

7. **Central quadrics.** The locus S of a second degree equation in Euclidean $(x_1, \cdots, x_n)$-space will be called a *central quadric* if it is not vacuous and is symmetric in a point P not on it, called a *center*. For example, if $n = 2$ the central quadrics (now conics) are of the following types: circle, ellipse, hyperbola, pair of parallel straight lines. In general a quadric surface is symmetric and has only one center. In any case, a particular center will be chosen and called *the center*.

[21] Cf. Mayer, pp. 158–159.

[22] By Theorem 5.1, the projections on the $y_{n+1}, \cdots, y_m$ directions are the same for all the curves, and as the projections on the $y_1, \cdots, y_n$ directions are zero, the vectors are all the same.

[23] Eisenhart, p. 75; Mayer, p. 159.

[24] Cf. Eisenhart, p. 113.

[25] This part was originally presented to the Society under its own title.

In the earlier part of this paper we had occasion to consider the value of an arbitrary real quadratic form $a_{ij}x_ix_j$ on the locus $x_ix_i = 1$. This suggests considering the value of x_ix_i on the locus $a_{ij}x_ix_j = 1$. Our principal result is that if for a central quadric in n-space the lengths of the diameters, as functions of parameters determining the direction, have critical points in only a finite number of directions, the i-th in magnitude is at a non-degenerate critical point of index $i - 1$.

8. **Preliminaries.** If the center is at the origin, x_ix_i is one-fourth the square of the diameter with end points (x) and $(-x)$. As we prefer to consider the diameter itself, we begin with the following lemma.

LEMMA. *Suppose $f(x_1, \cdots, x_n)$ is of class C^2 neighboring (x^0) in real n-space, $g(x) = [f(x)]^2$, and $f(x^0) > 0$. Then if either f or g has a non-degenerate critical point at (x^0), the other has one of the same index.*

Proof. Since $\partial g/\partial x_i = (2f)(\partial f/\partial x_i)$, if either has a critical point so has the other. Differentiating again and setting $x_i = x_i^0$ $(i = 1, \cdots, n)$ we have

$$\frac{\partial^2 g(x^0)}{\partial x_i \partial x_j} = 2f(x^0)\frac{\partial^2 f(x^0)}{\partial x_i \partial x_j},$$

since $\partial f(x^0)/\partial x_i = 0$. The conclusion now follows easily from the definition of index.[26] We shall apply the lemma with n replaced by $n - 1$.

Now let us consider the principal question. Suppose a central quadric S is given in n-space, $n > 1$, and we make a translation of axes so that the center (or a center, which will be designated hereafter as the center) becomes the origin. The new equation has no first degree terms and as in §2 we can make an orthogonal transformation,[27] giving us the new equation $c_ix_i^2 = d$. Since the origin is not on S, $d \neq 0$ and can be made unity. We can easily arrange that S has equation

$$(8.1) \qquad a_1x_1^2 + \cdots + a_kx_k^2 - b_{k+1}x_{k+1}^2 - \cdots - b_nx_n^2 = 1$$

with the a's all positive and the b's positive or zero. Here $k \geqq 1$ since S is not vacuous. Note that $c_1 = a_1, \cdots, c_k = a_k, c_{k+1} = -b_{k+1}, \cdots, c_n = -b_n$.

We now seek the critical points of

$$(8.2) \qquad z = x_ix_i$$

on the locus (8.1), as a function of $n - 1$ parameters. A line through any point on (8.1) directed towards the origin is easily seen not to be orthogonal to (8.1). Using this fact we easily prove that, as in a similar situation in §2, it is sufficient to find the critical points located on (8.1) of the function

$$(8.3) \qquad w = (x_ix_i)/(c_ix_i^2)$$

[26] If $[\partial^2 f(x^0)\partial x_i \partial x_j]\cdot x_ix_j = -y_1^2 - \cdots - y_k^2 + y_{k+1}^2 + \cdots + y_n^2$ under a non-singular linear transformation, the index is k.

[27] The determinant of the coefficients can be made $+1$ if we like, so that the transformation is a "rigid motion".

of n independent variables. Upon differentiation we find that they are the points satisfying (8.1) and

$$(8.4) \qquad x_s = (x_i x_i) c_s x_s \qquad (s = 1, \cdots, n; \text{ do not sum } s).$$

Since $x_i x_i \neq 0$ on (8.1), at any critical point at least one $x_s \neq 0$, hence by (8.2) and (8.4) $z = 1/c_s$. Since $c_{k+1}, \cdots, c_n$ are not positive, the number of distinct critical values of z is seen to be at most k, hence finite.

9. **The theorem.** A diameter in a direction where the length has a critical point will be called a *critical diameter*.

THEOREM. *If a central quadric in n-space, $n > 1$, has only a finite number of critical diameters, they are mutually orthogonal and the i-th in length is at a non-degenerate critical point of index $i - 1$.*

Proof. Suppose, in (8.1), $a_1 = a_2$. Then any point $(x_1, x_2, 0, \cdots, 0)$ such that $x_1^2 + x_2^2 = 1/a_1 = 1/a_2 = 1/c_1 = 1/c_2$ would satisfy (8.1) and (8.4), hence determine a critical diameter. The number of critical diameters would then be infinite, contrary to hypothesis. We infer that the a's are all distinct.

In §8 we saw that if $x_s \neq 0$ at a critical point, $z = 1/c_s$ at the point. Hence $s \leqq k$, and since $c_1, \cdots, c_k$ are all distinct, at most one x is different from zero. Hence the critical diameters are mutually orthogonal. Since the x's cannot be all zero, we infer that there are just k critical diameters, one along each of the first k coördinate axes.

Consider the solutions of (8.1), (8.4) with $x_1 \neq 0$. They are

$$(\pm 1/(a_1)^{\frac{1}{2}}, 0, \cdots, 0).$$

We choose either sign. Neighboring the point chosen, $x_2, \cdots, x_n$ can be taken as independent variables for (8.1). Substituting from (8.1) into (8.2) we obtain

$$z = \frac{1}{a_1} + \left(1 - \frac{a_2}{a_1}\right)x_2^2 + \cdots + \left(1 - \frac{a_k}{a_1}\right)x_k^2 + \left(1 + \frac{b_{k+1}}{a_1}\right)x_{k+1}^2 + \cdots + \left(1 + \frac{b_n}{a_1}\right)x_n^2.$$

Since none of the coefficients is zero, the critical point is non-degenerate and of index equal to the number of a's larger than a_1, hence to the number of $(1/a)$'s smaller than $(1/a_1)$. The critical value is $1/a_1$. Since a similar result is obtained with x_1 replaced by each of $x_2, \cdots, x_k$, we infer the truth of the theorem.

REMARK. *If S is a central quadric not a surface of revolution, the hypotheses of the theorem are satisfied.* For we have seen that the distinctness of the a's implies that the critical diameters are mutually orthogonal, hence finite in number.

On the other hand, a surface of revolution may satisfy the hypotheses of the theorem, for example, any surface for which the a's are distinct but two of the b's are equal.

COLUMBIA UNIVERSITY.

THE CHARACTERISTIC ROOTS OF A MATRIX

By W. V. Parker

1. **Introduction.** If A is a square matrix of order n and I is the unit matrix, the equation obtained by equating to zero the determinant $| A - \lambda I |$ is called the *characteristic equation* of A. The roots of this equation are called the *characteristic roots* of A.

If A is a matrix of a particular type, certain definite statements may be made concerning the nature of its characteristic roots. For example, if A is Hermitian its characteristic roots are all real. While it is not possible to make any definite statement about the nature of the characteristic roots for the general matrix, several authors have given upper limits to the roots. The first of these limits seems to have been given by Bendixson[1] in 1900. He obtained upper limits for the real and imaginary parts of the characteristic roots of a *real* matrix. In a letter to Bendixson in 1902, Hirsch[2] extended these results to include the case when the elements of A may be complex numbers. Hirsch obtained an upper limit for the characteristic roots as well as for their real and imaginary parts. A limit was also given by Bromwich[3] in 1904. In 1930, Browne[4] obtained limits which do not exceed those previously found and are in general less.

In the present note it is shown that the limit for the characteristic roots can generally be determined to be less than the one given by Browne. A lower limit for the characteristic root of greatest absolute value and an upper limit for the characteristic root of least absolute value for an Hermitian matrix are also found.

Let A' and $\bar{A}$ denote the transpose and conjugate, respectively, of the square matrix A and write

$$B = \frac{A + \bar{A}'}{2}, \qquad C = \frac{A - \bar{A}'}{2i}.$$

It is evident that B and C are Hermitian; that is, $B = \bar{B}'$ and $C = \bar{C}'$. A theorem given by Browne may be stated as follows:

Browne's Theorem. *If R_i, R_i', and R_i'' are the sums of the absolute values of the elements in the i-th row of the matrices A, B, and C, respectively, and if T_i*

Received March 10, 1937.

[1] Bendixson, *Sur les racines d'une équation fondamentale*, Acta Mathematica, vol. 25 (1902), pp. 359–365.

[2] Hirsch, Acta Mathematica, vol. 25 (1902), pp. 367–370.

[3] Bromwich, *On the roots of the characteristic equation of a linear substitution*, Acta Mathematica, vol. 30 (1906), pp. 295–304.

[4] Browne, *The characteristic roots of a matrix*, Bulletin of the American Mathematical Society, vol. 36 (1930), pp. 705–710.

is the sum of the absolute values of the elements in the i-th column of A, and if R, R', R'', and T are the greatest of the R_i, R'_i, R''_i, and T_i, respectively, then for any characteristic root, $\lambda = \alpha + i\beta$, of A we have

$$|\lambda| \leqq \frac{R+T}{2}, \qquad |\alpha| \leqq R', \qquad |\beta| \leqq R''.$$

2. **An upper limit to the characteristic roots of** A. If $\lambda = \alpha + i\beta$ is a characteristic root of a matrix $A = (a_{rs})$ of order n, there exists a set of numbers $(x_1, x_2, \cdots, x_n)$ such that $\sum_{r=1}^{n} x_r \bar{x}_r = 1$, which satisfy the relations

$$\lambda x_r = \sum_{s=1}^{n} a_{rs} x_s \qquad (r = 1, 2, \cdots, n). \tag{1}$$

If we multiply the r-th equation in (1) by $\bar{x}_r$ and sum as to r, we get

$$\lambda = \sum_{r,s=1}^{n} a_{rs} \bar{x}_r x_s . \tag{2}$$

If in (2) we take conjugates on both sides and interchange the subscripts, we get

$$\bar{\lambda} = \sum_{r,s=1}^{n} \bar{a}_{sr} \bar{x}_r x_s . \tag{3}$$

From (2) and (3) by addition and subtraction we get

$$\alpha = \sum_{r,s=1}^{n} b_{rs} \bar{x}_r x_s , \tag{4}$$

$$\beta = \sum_{r,s=1}^{n} c_{rs} \bar{x}_r x_s . \tag{5}$$

From the relations (2), (4) and (5) we determine upper limits for $|\lambda|$, $|\alpha|$ and $|\beta|$. Since these relations are identical in form, it is sufficient to carry the computation through for one of them. We shall write

$$R_r = \sum_{s=1}^{n} |a_{rs}|, \qquad T_s = \sum_{r=1}^{n} |a_{rs}|, \qquad 2S_r = R_r + T_r,$$

and let $\xi_r = |x_r|$ so that $\sum_{r=1}^{n} \xi_r^2 = 1$ and $\xi_r \xi_s \leqq \frac{1}{2}(\xi_r^2 + \xi_s^2)$. If we take absolute values in (2), we get

$$|\lambda| \leqq \sum_{r,s=1}^{n} |a_{rs}| \xi_r \xi_s \leqq \frac{1}{2} \sum_{r,s=1}^{n} |a_{rs}|(\xi_r^2 + \xi_s^2) = \frac{1}{2} \sum_{r=1}^{n} R_r \xi_r^2 + \frac{1}{2} \sum_{s=1}^{n} T_s \xi_s^2$$

$$= \sum_{r=1}^{n} S_r \xi_r^2 \leqq S \sum_{r=1}^{n} \xi_r^2 = S,$$

where S is the greatest of the S_r. Similar inequalities hold for $|\alpha|$ and $|\beta|$ and we have the following theorem.

THEOREM 1. *If A is any square matrix, and if $2S_r$, $2S_r'$, $2S_r''$ are the sums of the absolute values of the elements in the r-th row and the absolute values of the elements in the r-th column of A, B, and C, respectively, and if S, S', S'' are the greatest of the S_r, S_r', S_r'', respectively, then for any characteristic root, $\lambda = \alpha + i\beta$, of A we have*

$$|\lambda| \leqq S, \qquad |\alpha| \leqq S', \qquad |\beta| \leqq S''.$$

The latter two limits are those given by Browne but the limit for $|\lambda|$ is generally less than his.

3. **The characteristic roots of $A\bar{A}'$.** Browne[5] has shown that if λ is a characteristic root of a square matrix A and if G is the greatest and s the smallest of the (real and not negative) characteristic roots of $A\bar{A}'$ then $0 \leqq s \leqq \lambda\bar{\lambda} \leqq G$.

Let U_r be a square matrix of order n whose elements are determined by the following conditions. The elements of the r-th column (r a definite number) are

$$\bar{a}_{r1}\sigma_r^{-\frac{1}{2}},\ \bar{a}_{r2}\sigma_r^{-\frac{1}{2}},\ \cdots,\ \bar{a}_{rn}\sigma_r^{-\frac{1}{2}},$$

where

$$\sigma_r = \sum_{t=1}^{n} a_{rt}\bar{a}_{rt} > 0.\text{[6]} \tag{6}$$

The elements of the j-th column ($j \neq r$) are

$$\bar{x}_{j1},\ \bar{x}_{j2},\ \cdots,\ \bar{x}_{jn},$$

where

$$\begin{aligned} &\sum_{t=1}^{n} a_{rt}\bar{x}_{jt} = 0 \quad (j \neq r), \\ &\sum_{t=1}^{n} x_{it}\bar{x}_{jt} = \delta_{ij} \quad (i, j \neq r). \end{aligned} \tag{7}$$

The matrix U_r thus determined is unitary, that is, $\bar{U}_r'U_r = I$. If $P_r = AU_r$, the elements of the r-th row of P_r are

$$p_{rj} = \sum_{t=1}^{n} a_{rt}\bar{x}_{jt} = 0 \quad (j \neq r),$$

$$p_{rr} = \sum_{t=1}^{n} a_{rt}\bar{a}_{rt}\sigma_r^{-\frac{1}{2}} = \sigma_r^{\frac{1}{2}}.$$

It is evident, therefore, that $\sigma_r^{\frac{1}{2}}$ is a characteristic root of the matrix $P_r = AU_r$.

[5] Browne, *The characteristic equation of a matrix*, Bulletin of the American Mathematical Society, vol. 34 (1928), pp. 363–368.

[6] It is assumed here that not all elements of the r-th row of A are zero. If all elements of any one row are zero, then zero is a characteristic root of A and the theorems as stated here are still true.

In a similar way we may construct a unitary matrix V_s such that $Q_s = V_s A$ has the characteristic root $\tau_s^{\frac{1}{2}}$ where

$$\tau_s = \sum_{t=1}^{n} \bar{a}_{ts} a_{ts} . \tag{8}$$

By taking products we see that

$$P_r \bar{P}'_r = A U_r \bar{U}'_r \bar{A}' = A\bar{A}',$$
$$\bar{Q}'_s Q_s = \bar{A}' \bar{V}'_s V_s A = \bar{A}' A.$$

Hence the characteristic roots of $P_r \bar{P}'_r$ and $\bar{Q}'_s Q_s$ are identical with those of $A\bar{A}'$. These roots are real and not negative. If G is the greatest and s the smallest of these roots, we have, as shown by Browne

$$s \leqq \sigma_r \leqq G \text{ and } s \leqq \tau_s \leqq G.$$

Hence we have the following theorem.

THEOREM 2. *If $\sigma_r(\tau_r)$ is the sum of the squares of the absolute values of the elements of the r-th row (column) of a square matrix A, and if $\sigma(\tau)$ is the greatest and $\sigma'(\tau')$ the smallest of the $\sigma_r(\tau_r)$, then $A\bar{A}'$ has a real positive (or zero) characteristic root not less than the greater of σ and τ and a real positive (or zero) characteristic root not greater than the smaller of σ' and τ'.*

In particular, if A is Hermitian, its characteristic roots are all real and the characteristic roots of $A\bar{A}'$ are the squares of the characteristic roots of A. If the characteristic roots of $A\bar{A}'$ are $\lambda_1^2 \geqq \lambda_2^2 \geqq \cdots \geqq \lambda_n^2$, we have

$$\lambda_1^2 \geqq \sigma \geqq \sigma' \geqq \lambda_n^2 ,$$

and hence

$$| \lambda_1 | \geqq \sigma^{\frac{1}{2}} \geqq (\sigma')^{\frac{1}{2}} \geqq | \lambda_n |.$$

Therefore we have the following theorem.

THEOREM 3. *If σ_r is the sum of the squares of the absolute values of the elements of the r-th row of a Hermitian matrix A, and if σ is the greatest and σ' the least of the σ_r, then A has at least one characteristic root whose absolute value does not exceed $(\sigma')^{\frac{1}{2}}$ and at least one characteristic root whose absolute value is not less than $\sigma^{\frac{1}{2}}$.*

4. **A lower limit for the characteristic roots of** A. A matrix has the characteristic root zero if and only if it is singular. It is well known that if λ is a characteristic root of a non-singular matrix A, then $1/\lambda$ is a characteristic root of A^{-1}, the inverse of A. If $| 1/\lambda | \leqq L$, then $| \lambda | \geqq 1/L$ and hence an upper limit for the characteristic roots of A^{-1} determines a lower limit for the characteristic roots of A.

LOUISIANA STATE UNIVERSITY.

COMPLETELY MONOTONE FUNCTIONS OF THE LAPLACE OPERATOR FOR TORUS AND SPHERE

By S. Bochner

In the present note we shall discuss some properties of the Laplace operator

$$\Delta g = -\frac{1}{(2\pi)^k}\left(\frac{\partial^2}{\partial x_1^2} + \cdots + \frac{\partial^2}{\partial x_k^2}\right)g \tag{1}$$

on the *torus*

$$-\tfrac{1}{2} \leqq x_\kappa < \tfrac{1}{2} \qquad (\kappa = 1, \cdots, k), \tag{2}$$

that is, for functions having the period 1 in each variable, and corresponding properties for the Laplace-Beltrami operator on the sphere of positive constant curvature.

Part I. The torus

1. We introduce on the torus (2) the Hilbert space of functions of integrable square. It is the family of functions

$$f(x) \sim \sum \overset{\infty}{\underset{-\infty}{\cdots}} \sum a_{n_1 \cdots n_k} \exp[2\pi i(n_1 x_1 + \cdots + n_k x_k)] \tag{3}$$

for which

$$\sum \overset{\infty}{\underset{-\infty}{\cdots}} \sum |a_{n_1 \cdots n_k}|^2 < \infty. \tag{4}$$

At the outset, the operator (1) is defined only for functions which are differentiable twice. As such it is a positive semi-definite Hermitian operator with characteristic functions

$$\exp[2\pi i(n_1 x_1 + \cdots + n_k x_k)] \tag{5}$$

belonging to the characteristic values

$$n_1^2 + \cdots + n_k^2. \tag{6}$$

In accordance with a general theorem referring to the nature of the Laplace-Beltrami operator on a compact Riemann space, our initial operator has a unique self-adjoint (hyper-maximal) closure with the same spectrum.[1] In what follows we shall be interested in this closure only and we shall denote it by Δg. The spectrum of Δg is non-negative. Let $\varphi(\rho)$ be an arbitrary real *continuous*

Received May 4, 1937.

[1] Cf. S. Bochner, *Analytic mapping of compact Riemann spaces into Euclidean space*, this Journal, vol. 3 (1937), pp. 339–354.

function in the half-line $0 \leqq \rho < \infty$. According to a general theorem on functions of operators[2] we can form the operator $\varphi(\Delta)g$. It is again self-adjoint and, in our case, it is uniquely determined by the fact that it again has the characteristic functions (5), the corresponding characteristic values being

$$\varphi(n_1^2 + \cdots + n_k^2).$$

If $\varphi(\Delta)f$ is defined for an element (3), then its value is

$$\sum_{-\infty}^{\infty} \varphi(n_1^2 + \cdots + n_k^2) a_{n_1 \cdots n_k} \exp [2\pi i(n_1 x_1 + \cdots + n_k x_k)], \tag{7}$$

and it is defined for an element (3) if and only if (7) is again an element of the Hilbert space, that is, if

$$\sum_{-\infty}^{\infty} | \varphi(n_1^2 + \cdots + n_k^2) |^2 \cdot | a_{n_1 \cdots n_k} |^2 < \infty.$$

For example, if

$$\varphi(\rho) = \frac{1}{\rho + c}, \qquad c > 0, \tag{8}$$

then (7) assumes the form

$$\sum_{-\infty}^{\infty} \frac{a_{n_1 \cdots n_k}}{n_1^2 + \cdots + n_k^2 + c} \exp [2\pi i(n_1 x_1 + \cdots + n_k x_k)],$$

and in this case the operator $\varphi(\Delta)f$ is the inverse of the operator $(\Delta + c)g$; that is, it represents the solution of the equation

$$(\Delta + c)g = f. \tag{9}$$

Purely formally the expression (7) can be written as an integral operator. Introduce the Green's function

$$G(x) = \sum_{-\infty}^{\infty} \varphi(n_1^2 + \cdots + n_k^2) \exp [2\pi i(n_1 x_1 + \cdots + n_k x_k)], \tag{10}$$

then

$$\varphi(\Delta)f = \int \cdots \int_{x_k = -\frac{1}{2}}^{+\frac{1}{2}} G(x_1 - \xi_1, \cdots, x_k - \xi_k) f(\xi_1, \cdots, \xi_k) \, d\xi_1 \cdots d\xi_k. \tag{11}$$

Our aim is to discuss a class of functions $\varphi(\rho)$ for which this representation is valid.

2. We assume that $\varphi(\rho)$ is completely monotone in $0 \leqq \rho < \infty$. This means that $\varphi(\rho)$ has derivatives of all orders in $0 \leqq \rho < \infty$ and that

$$(-1)^n \frac{d^n \varphi(\rho)}{d\rho^n} \geqq 0 \qquad (n = 0, 1, 2, \cdots).$$

[2] M. H. Stone, *Linear Operations in Hilbert Space*, 1932, Chapters VI, VII.

By an important theorem of S. Bernstein and Widder[3] any function which is completely monotone in a half-line $\rho_0 < \rho < \infty$ can be represented uniquely as a Laplace integral

$$\varphi(\rho) = \int_0^\infty e^{-\rho t}\, d\alpha(t) \tag{12}$$

in which the function $\alpha(t)$ is monotonely non-decreasing. Since, by our assumption, $\varphi(\rho)$ is bounded to the right of $\rho = 0$, we have

$$\int_0^\infty d\alpha(t) < \infty; \tag{13}$$

this assumption will be somewhat relaxed later on. Another restriction which will be required throughout is the condition

$$\alpha(+0) = \alpha(0); \tag{14}$$

thus the function $\alpha(t)$ shall have no discontinuity for $t = 0$. The rôle of this restriction is obvious; it excludes the function $\varphi(\rho) = 1$ for which no Green's function exists. As a consequence of this restriction we have the limit relation

$$\varphi(\rho) = \lim_{a \to 0} \varphi_a(\rho), \tag{15}$$

where

$$\varphi_a(\rho) = \int_a^\infty e^{-\rho t}\, d\alpha(t), \qquad a > 0. \tag{16}$$

Obviously

$$\varphi_a(\rho) = O(e^{-\rho a}), \qquad \rho \to \infty. \tag{17}$$

Postponing questions of convergence, we form the integral

$$\begin{aligned} &H(x_1, \cdots, x_k) \\ &= \int_{-\infty}^{\infty} \cdots \int \varphi(n_1^2 + \cdots + n_k^2) \exp\,[2\pi i(n_1 x_1 + \cdots + n_k x_k)]\, dn_1 \cdots dn_k, \end{aligned} \tag{18}$$

and we observe that, by Poisson's summation formula,[4] it is connected with the function (10) by the relation

$$G(x_1, \cdots, x_k) = \sum_{m_k=-\infty}^{\infty} \cdots \sum H(x_1 + m_1, \cdots, x_k + m_k). \tag{19}$$

Substituting in (18)

$$\varphi(n_1^2 + \cdots + n_k^2) = \int_0^\infty \exp\,[-(n_1^2 + \cdots + n_k^2)t]\, d\alpha(t),$$

[3] D. V. Widder, *Necessary and sufficient conditions for the representation of a function as a Laplace integral*, Trans. Amer. Math. Soc., vol. 33 (1931), pp. 851–892.

[4] Cf. S. Bochner, *Vorlesungen über Fouriersche Integrale*, 1932, pp. 33–38, 203–205.

inverting the integrations and making use of the formula

$$\int_{-\infty}^{\infty} \exp[-(n^2t - 2\pi i n x)]dn = (\pi t)^{-\frac{1}{2}} \exp\left(-\frac{\pi^2 x^2}{t}\right),$$

we obtain

$$H(x_1, \cdots, x_k) = \pi^{-\frac{1}{2}k} \int_0^{\infty} \exp\left[-\frac{\pi^2(x_1^2 + \cdots + x_k^2)}{t}\right] \cdot t^{-\frac{1}{2}k} d\alpha(t). \tag{20}$$

Hence, introducing the function of one variable

$$H(r) = \pi^{-\frac{1}{2}k} \int_0^{\infty} \exp\left[-\frac{\pi^2 r^2}{t}\right] \cdot t^{-\frac{1}{2}k} d\alpha(t), \tag{21}$$

we have

$$H(x_1, \cdots, x_k) = H[(x_1^2 + \cdots + x_k^2)^{\frac{1}{2}}].$$

The function $H(r)$ which we consider for $0 < r < \infty$ is positive and non-increasing, by (20), and the function $G(x)$ is positive, by (19).

In order to justify these relations we first replace the function $\varphi(\rho)$ by the truncated function (16) and we denote the corresponding functions (10), (18), (21) by $G_a(x)$, $H_a(x)$, $H_a(r)$, and the corresponding operator $\varphi(\Delta)f$ by $\varphi_a(\Delta)f$. Because of (17), the expressions (10), (18), (21) converge absolutely, relation (19) is valid, and, the function $G_a(x)$ being bounded, the right side of (11) has a meaning, and its value is (7), for any square-integrable function $f(x)$. Thus relation (11) holds.

We now assume that a tends decreasingly to 0. Initially, the integral (18) and the sum (10) have no meaning for the function $\varphi(\rho)$ itself, but we *define* them by the limit-relations

$$H(x) = \lim_{a \to 0} H_a(x), \qquad G(x) = \lim_{a \to 0} G_a(x). \tag{22}$$

Since the functions $H_a(x)$, $G_a(x)$ increase when a decreases, the limit relations have a meaning (although the resulting functions might have values $+\infty$), and relation (19) holds. We shall now use relation (15). In the first place it justifies relation (20), thus proving that $H(r)$ is everywhere finite. But we have still to show the finiteness of (10) and the validity of (11). As long as (18) is absolutely convergent it admits the inversion formula,

$$\begin{aligned} &\varphi(n_1^2 + \cdots + n_k^2) \\ &= \int_{-\infty}^{\infty} \cdots \int H(x_1, \cdots, x_k) \exp[-2\pi i(x_1 n_1 + \cdots + x_k n_k)] dx_1 \cdots dx_k. \end{aligned} \tag{23}$$

In particular,

$$\varphi_a(0) = \int_{-\infty}^{\infty} \cdots \int H_a(x_1, \cdots, x_k)\, dx_1 \cdots dx_k.$$

Since $H_a(x_1, \cdots, x_k)$ and $\varphi_a(0)$ tend increasingly to their limit-values, we obtain

$$\varphi(0) = \int_{-\infty}^{\infty} \cdots \int H(x_1, \cdots, x_k)\, dx_1 \cdots dx_k.$$[5]

Thus, the integral on the right is bounded. Since $H(r)$ decreases as r increases, we readily conclude that the sum (19), but for the term $H(x_1, \cdots, x_k)$, converges uniformly in (2). Thus, in (2), $G(x_1, \cdots, x_k)$ differs from $H(x_1, \cdots, x_k)$ by a bounded function only and, in particular, is finite. Since, by (10),

$$\varphi_a(0) = \int_{\xi_\kappa = -\frac{1}{2}}^{+\frac{1}{2}} \cdots \int G_a(\xi_1, \cdots, \xi_k)\, d\xi_1 \cdots d\xi_k,$$

the integral on the right is finite also for the limit-function $G(\xi_1, \cdots, \xi_k)$. It is now easy to give a meaning to the integral (11) and to show that it has the value (7). As a consequence of (14) the function $\varphi(\rho)$ tends to 0 as $\rho \to \infty$ and the value (7) of $\varphi(\Delta)f$ shows that all our functions $\varphi(\Delta)$ of Δ are completely continuous operators.

3. We shall next discuss the order of infinity of $G(x_1, \cdots, x_k)$ in the neighborhood of the origin in terms of the magnitude of $\varphi(\rho)$ as $\rho \to \infty$. In the statements the function $G(x_1, \cdots, x_k)$ will be replaced by the function $H(r)$.

I. *In order that $H(r)$ be bounded (in the neighborhood of $r = 0$) it is necessary and sufficient that for some (and therefore every) $a > 0$*

$$\int_a^{\infty} \varphi(\rho)\rho^{\frac{1}{2}k-1}\, d\rho < \infty. \tag{24}$$

In fact, all occurring functions being non-negative, we may conclude

$$\begin{aligned}\int_0^{\infty} \varphi(\rho)\rho^{\frac{1}{2}k-1}\, d\rho &= \int_0^{\infty}\int_0^{\infty} e^{-\rho t}\rho^{\frac{1}{2}k-1}\, d\rho\, d\alpha(t) = \Gamma\left(\frac{k}{2}\right)\int_0^{\infty} t^{-\frac{1}{2}k}\, d\alpha(t)\\ &= \Gamma\left(\frac{k}{2}\right)\pi^{\frac{1}{2}k}\cdot\lim_{r\to 0} H(r).\end{aligned} \tag{25}$$

To simplify writing we shall put

$$K(r) = \int_0^{\infty} e^{-r/t}\, t^{-\frac{1}{2}k}\, d\alpha(t) = \pi^{\frac{1}{2}k} H\left(\frac{r^{\frac{1}{2}}}{\pi}\right).$$

Similar to (25) we obtain, for

$$0 < \lambda < \tfrac{1}{2}k,$$

[5] See also S. Bochner, *Monotone Funktionen, Stieltjesche Integrale und harmonische Analyse*, Math. Annalen, vol. 108 (1933), pp. 399–408.

putting $\mu = \frac{1}{2}k - \lambda$, the relation

$$\frac{1}{\Gamma(\lambda)} \int_0^\infty \varphi(\rho)\rho^{\lambda-1} d\rho = \frac{1}{\Gamma(\mu)} \int_0^\infty K(r) r^{\mu-1} dr.$$

Hence, we have

II. *In order that for some (and therefore every)* $a > 0$

$$\int_0^a H(r) r^{k-2\lambda-1} dr < \infty, \tag{26}$$

it is necessary and sufficient that for some (and therefore every) $b > 0$

$$\int_b^\infty \varphi(\rho)\rho^{\lambda-1} d\rho < \infty. \tag{27}$$

Finally, we have

III. *If* $L(\xi)$ *is any function in* $0 < \xi < \infty$ *such that, for every* $c > 0$,

$$\lim \frac{L(c\xi)}{L(\xi)} = 1 \quad \textit{as} \quad \xi \to 0 \quad \textit{and} \quad \xi \to \infty, \tag{28}$$

then the asymptotic relation

$$\varphi(\rho) \sim \frac{A}{\rho^\lambda} L\left(\frac{1}{\rho}\right), \qquad \rho \to \infty, \tag{29}$$

implies an asymptotic relation

$$H(r) \sim \frac{B}{r^{k-2\lambda}} L(r^2), \qquad r \to 0, \tag{30}$$

and vice versa.

This follows from the following theorem.[6]

If $T(t)$ *is non-decreasing,* $T(+0) = T(0)$, *and* $\int_0^\infty e^{-st}\, dT(t)$ *is finite for* $s > 0$, *then, for* $\sigma > 0$, *the relations*

$$\int_0^\infty e^{-st} dT(t) \sim s^{-\sigma} L\left(\frac{1}{s}\right) \qquad \begin{matrix} \textit{as } s \to 0 \\ \textit{as } s \to \infty \end{matrix} \tag{31}$$

and

$$T(x) \sim \frac{x^\sigma L(x)}{\Gamma(\sigma + 1)} \qquad \begin{matrix} \textit{as } x \to \infty \\ \textit{as } x \to 0 \end{matrix} \tag{32}$$

are equivalent.

We shall verify, for instance, that (29) implies (30) or, what is the same, that (29) implies

$$K(r) \sim \frac{C}{r^\mu} L(r), \qquad r \to 0.$$

[6] J. Karamata, *Neuer Beweis und Verallgemeinerung der Tauberschen Sätze, welche die Laplacesche und Stieltjesche Transformation betreffen*, Crelle's Journal, vol. 146 (1931), pp. 27–39.

By assumption

$$\int_0^\infty e^{-st}\,d\alpha(t) \sim \frac{A}{s^\lambda} L\left(\frac{1}{s}\right), \qquad s \to \infty.$$

Consequently, by the quoted theorem,

$$\alpha(t) \sim \frac{At^\lambda L(t)}{\Gamma(\lambda + 1)}, \qquad t \to 0,$$

or

$$\alpha\left(\frac{1}{t}\right) \sim \frac{A}{\Gamma(\lambda + 1)} \cdot \frac{1}{t^\lambda} L\left(\frac{1}{t}\right), \qquad t \to \infty.$$

Defining the function $T(t)$ by the relation

$$dT(t) = t^{\frac{1}{2}k} d\left[-\alpha\left(\frac{1}{t}\right)\right]$$

we easily obtain

$$T(x) \sim \frac{A}{\Gamma(\lambda)} \cdot \frac{x^\mu}{\mu} L\left(\frac{1}{x}\right), \qquad x \to \infty$$

But

$$K(s) = \int_0^\infty e^{-s/t}\, t^{-\frac{1}{2}k}\, d\alpha(t) = \int_0^\infty e^{-st}\, dT(t),$$

and therefore, since (32) implies (31), we obtain, replacing $L(x)$ by $L(x^{-1})$,

$$K(s) \sim \frac{A\Gamma(\mu)}{\Gamma(\lambda)} \cdot \frac{1}{s^\mu} L(s), \qquad s \to 0.$$

4. Finally we shall show that $G(x_1, \cdots, x_k)$ is analytic at all points of the torus except the origin and that for an analytic function $f(x_1, \cdots, x_k)$ the transformed function $\varphi(\Delta)f$ is again analytic.

If $x_1^0, \cdots, x_k^0$ is any fixed point on the torus different from $(0, \cdots, 0)$, then the function $H(x)$ is analytic in some neighborhood of this point. In fact, given an $\epsilon > 0$ there exists a $\delta > 0$ such that

$$|x_1|^2 + \cdots + |x_k|^2 \geqq |x_1^0|^2 + \cdots + |x_k^0|^2 - \delta$$

for all *complex* values of the neighborhood

$$|x_1 - x_1^0|^2 + \cdots + |x_k - x_k^0|^2 < \epsilon.$$

Consequently, the integral (20) and the partial derivatives

$$(33) \qquad \frac{\partial H}{\partial x_\kappa} = \pi^{-\frac{1}{2}k} \int_0^\infty \left(-\frac{2\pi^2 x_\kappa}{t}\right) \exp\left(-\frac{\pi^2}{t}(x_1^2 + \cdots + x_k^2)\right) t^{-\frac{1}{2}k}\, d\alpha(t)$$

exist in a *complex* neighborhood, thus proving the analyticity of $H(x_1, \cdots, x_k)$ at all points but the origin. Also, the argument in §2 shows easily that the

series (19) converges uniformly in our complex neighborhood; the result is that $G(x_1, \cdots, x_k)$ is analytic at all points of the torus except the origin and that in the neighborhood of the origin it differs from the function $H(x_1, \cdots, x_k)$ by an analytic function. Furthermore, in dealing with the singularity of $H(x_1, \cdots, x_k)$ at the origin we may assume that $\varphi(\rho)$ has the form

$$\varphi^*(\rho) = \int_0^a e^{-\rho t}\, d\alpha(t). \tag{34}$$

This follows from the fact that the difference function $\varphi_a(\rho) = \varphi(\rho) - \varphi^*(\rho)$ (see (16)) has the order of magnitude

$$\varphi_a(n_1^2 + \cdots + n_k^2) = O\,(\exp[-4\pi(n_1^2 + \cdots + n_k^2)])$$

for $a > 0$ sufficiently large, thus making the corresponding function $G_a(x)$ analytic by crude absolute and uniform convergence of series (10).

We now assume that $f(x_1, \cdots, x_k)$ is analytic on the torus. If x runs over a complex neighborhood of a fixed point and ξ over a set S of the torus such that $G(x - \xi)$ is analytic in the neighborhood of the fixed point, uniformly in ξ from S, then the integral

$$\int \cdots \int_S G(x - \xi) f(\xi)\, d\xi_1 \cdots d\xi_k$$

is analytic at the given point x (even if $f(\xi)$ is not analytic). Hence it is sufficient to show that

$$g(x) = \int \cdots \int_{S_r} H(x - \xi) f(\xi)\, d\xi_1 \cdots d\xi_k \tag{35}$$

is analytic at the origin if S_r is some sphere $\xi_1^2 + \cdots + \xi_k^2 \leqq r^2$ and

$$H(x) = \int_0^a \exp\left[-\frac{c}{t}(x_1^2 + \cdots + x_k^2)\right] \cdot t^{-\frac{1}{2}k}\, d\alpha(t),$$

$a > 0$, $c > 0$. Following E. E. Levi[7] we replace in (35) the coördinates $\xi_1, \cdots, \xi_k$ by a type of polar coördinates originating at the variable point x which is interior to S_r. The new coördinates consist of angular variables $\theta_1, \cdots, \theta_{k-1}$ and a radial distance ρ, $0 < \rho < 1$. The angular variables do *not* depend on the variable point x. They are fixed variables on the surface of the sphere S_r, such that the rectangular coördinates $\eta_1, \cdots, \eta_k$ of the surface $\eta_1^2 + \cdots + \eta_k^2 = r^2$ are fixed functions of $\theta_1, \cdots, \theta_{k-1}$. But the radial distance ρ does depend on x and is defined in the following way. Let η be any point on S_r; if ξ lies on the segment joining x and η, then

$$\xi_\kappa = x_\kappa + \rho(\eta_\kappa - x_\kappa) \qquad (\kappa = 1, \cdots, k),$$

[7] Cf. E. Hopf, *Über den funktionalen, insbesondere den analytischen Charakter der Lösungen elliptischer Differentialgleichungen zweiter Ordnung*, Math. Zeitschrift, vol. 34 (1931), p. 224.

the quantity ρ being the length of the segment (x, ξ) divided by the length of the segment (x, η). If we introduce these coördinates in (35), then the factor of $H(x - \xi)$ is analytic in a complex neighborhood of $x = 0$, uniformly in the variables $\theta_1, \cdots, \theta_{k-1}, \rho$, over which we integrate. Also, what is very important, the volume element contains the factor $\rho^{k-1}\,d\rho$. As for $H(x - \xi)$ itself, we can write it in the form

$$(36)\quad H(x-\xi) = \int_0^a \exp\left[-\frac{c\rho^2}{t}\left((x_1-\eta_1)^2 + \cdots + (x_k-\eta_k)^2\right)\right] t^{-\frac{1}{2}k}\,d\alpha(t).$$

Since $\eta_1^2 + \cdots + \eta_k^2 = r^2$, there exists a fixed *complex* neighborhood N of the point $x = 0$, such that in this neighborhood

$$|(x_1-\eta_1)^2 + \cdots + (x_k-\eta_k)^2| \geqq \frac{r^2}{2}.$$

Hence (36) is dominated by

$$\int_0^a \exp\left[-\frac{c\rho^2}{2t}r^2\right]\cdot t^{-\frac{1}{2}k}\,d\alpha(t).$$

Therefore (36) is analytic for $x \subset N$, uniformly in all $\theta_1, \cdots, \theta_{k-1}$ and in $\epsilon \leqq \rho \leqq r$, for any fixed $\epsilon > 0$. Hence the function

$$g_\epsilon(x) = \int_{S_r - S_\epsilon}\cdots\int H(x-\xi)f(\xi)\,d\xi_1\cdots d\xi_k$$

is analytic for $x \subset N$. But $g_\epsilon(x) - g(x)$ is dominated by

$$\int_0^\infty\int_0^a \exp\left[-\frac{c\rho^2}{2t}r^2\right] t^{-\frac{1}{2}k}\rho^{k-1}\,d\rho\,d\alpha(t),$$

and this is a finite dominant independent of x. The function $g(x)$, being the limit of boundedly convergent functions over a complex domain, is likewise analytic.

5. The function $H(x_1, \cdots, x_k)$ entering the sum (19) is a Green's function itself, namely, the Green's function of the operator $\varphi(\Delta)f$ not for the torus (2) but for the whole Euclidean space; compare its definition (18). For the special function

$$\varphi(\rho) = \frac{1}{\rho + c} = \int_0^\infty e^{-\rho t}e^{-ct}\,dt$$

we have

$$H(r) = \pi^{-\frac{1}{2}k}\int_0^\infty \exp\left[-\frac{\pi r^2}{t} - ct\right]\cdot t^{-\frac{1}{2}k}\,dt,$$

and therefore

$$H(r) = 2\pi^{1-k}(rc^{-\frac{1}{2}})^{1-\frac{1}{2}k}K_{1-\frac{1}{2}k}(2\pi rc^{\frac{1}{2}}),$$

where $K_\nu(z)$ is the Hankel function of imaginary argument.[8] For $k \geqq 3$, its expansion in the neighborhood of the origin, apart from a factor independent of r and c, starts with the term

$$\frac{1}{r^{k-2}}. \tag{37}$$

The other terms contain each a positive power of c or a product of a positive power of c with $\log c$. Thus, for $c \to 0$, we obtain the term (37) alone, and this is actually the Green's function for the operator $\Delta^{-1}f$ corresponding to $\varphi(\rho) = \rho^{-1}$. But the corresponding sum (19) tends to $+\infty$ as $c \to 0$, and so does formally the series (10) since its constant term $\varphi(0)$ tends to $+\infty$, if $\varphi(\rho) \to \rho^{-1}$. The reason is that on the torus the operator $\Delta^{-1}f$ does not exist for all functions of integrable square, since Δf has the proper function $f(x) = 1$ corresponding to the characteristic value 0. But the operator $\Delta^{-1}f$ exists for all functions which are orthogonal to this proper function, that is, for all functions $f(x)$ in whose expansion (3) the constant term vanishes. More generally, for this restricted class of functions $f(x)$ the operator $\varphi(\Delta)f$ will exist if $\varphi(\rho)$ is (bounded and) completely monotone in $1 \leqq \rho < \infty$. It will again be representable in the form (11) if in the sum (10) we omit the term corresponding to $n_1 = \cdots = n_k = 0$. Suppose in general that $\varphi(\rho)$ is completely monotone for $x \geqq b$, in which case $G(x)$ is defined as the sum

$$\sum_{n_1^2+\cdots+n_k^2 \geqq b^2} \cdots \sum \varphi(n_1^2 + \cdots + n_k^2) \exp\,[2\pi i(n_1x_1 + \cdots + n_kx_k)].$$

If a is any fixed number $\geqq b$, we put

$$\begin{aligned} \varphi(\rho) &= \int_0^a e^{-\rho t}\,d\alpha(t) + \int_a^\infty e^{-\rho t}\,d\alpha(t) \\ &= \varphi_1(\rho) + \varphi_2(\rho), \end{aligned}$$

and

$$\begin{aligned} G_1(x) &= \sum \mathop{\cdots}_{-\infty}^{\infty} \sum \varphi_1(n_1^2 + \cdots + n_k^2) \exp\,[2\pi i(n_1x_1 + \cdots + n_kx_k)], \\ G_2(x) &= -\sum_{n_1^2+\cdots+n_k^2<a^2} \cdots \sum \varphi_1(n_1^2 + \cdots + n_k^2) \exp\,[2\pi i(n_1x_1 + \cdots + n_kx_k)] \\ &\qquad + \sum_{n_1^2+\cdots+n_k^2 \geqq a^2} \cdots \sum \varphi_2(n_1^2 + \cdots + n_k^2) \exp\,[2\pi i(n_1x_1 + \cdots + n_kx_k)]. \end{aligned}$$

Obviously all previous considerations remain valid for the pair of functions $\varphi_1(\rho)$, $G_1(x)$. Also $\varphi_2(\rho) = O(e^{-\rho a})$ as $\rho \to \infty$, and therefore $G_2(x)$ is analytic and bounded on the torus. Consequently all previously stated properties remain in force for the functions $\varphi(\rho)$, $G(x)$ themselves.

[8] G. N. Watson, *Bessel Functions*, pp. 185–189.

Part II. The sphere

We shall now consider the Laplace-Beltrami operator on a k-dimensional space of constant positive curvature, and we shall assume that our space is given as a Euclidean sphere of radius 1,

$$\xi_1^2 + \cdots + \xi_{k+1}^2 = 1. \tag{38}$$

It is possible to prove analogues of our previous theorems for the most general completely monotone functions $\varphi(\rho)$ as before. But the formulas are more complicated and the argument is rather tedious. However, there are special classes of completely monotone functions for which special sets of formulas are available, and we shall restrict our attention to one special class of this kind.

6. We shall need a lemma on completely monotone functions.

If $\psi(\rho)$ is completely monotone in $\rho \geqq 0$,

$$\psi(\rho) = \int_0^\infty e^{-\rho t}\, d\alpha(t), \tag{39}$$

and $\chi(\rho)$ is the integral, vanishing at the origin, of a completely monotone function in $\rho \geqq 0$,

$$\chi'(\rho) = \int_0^\infty e^{-\rho t}\, d\gamma(t),$$

then the function $\varphi(\rho) = \psi(\chi(\rho))$ is again completely monotone in $\rho \geqq 0$.

The lemma can be proved directly from the definition by verifying inductively that the n-th derivative of $\varphi(\rho)$ has the sign of $(-1)^n$. However this procedure is rather cumbersome. A more elaborate but more illuminating proof runs as follows. We first conclude from the definition and the integral representation that sum, product and limit of completely monotone functions are again completely monotone. Therefore, since $\varphi(\rho)$ is a limit of finite sums of the form

$$\sum_n e^{-\chi(\rho)t_n}(\alpha(t_{n+1}) - \alpha(t_n)),$$

it is sufficient to prove that if $\chi(\rho)t_n$ is replaced by $\chi(\rho)$, $e^{-\chi(\rho)}$ is completely monotone. Approximating to $\chi'(\rho)$ by a finite sum

$$\sum_n e^{-\rho t_n}(\gamma(t_{n+1}) - \gamma(t_n))$$

we may further assume that $\chi'(\rho) = e^{-\rho t}$, in which case

$$\varphi(\rho) = \exp\left[\frac{e^{-\rho t} - 1}{t}\right] = e^{-1/t} \exp\left[\frac{e^{-\rho t}}{t}\right].$$

But the latter function is completely monotone, since

$$\exp\left[\frac{1}{t} e^{-\rho t}\right] = \sum_{n=0}^\infty \frac{e^{-n\rho t}}{n!\, t^n} = \int_0^\infty e^{-\rho u}\, d\beta(u).$$

In particular, if ν is any constant > 0, we may put

$$\chi(\rho) = (\rho + \nu^2)^{\frac{1}{2}} - \nu$$

since

$$2\chi'(\rho) = (\rho + \nu^2)^{-\frac{1}{2}} = \pi^{-\frac{1}{2}} \int_0^\infty e^{-\rho t} e^{-\nu^2 t} t^{-\frac{1}{2}} \, dt.$$

Hence, if $\psi(\rho)$ is any completely monotone function in $\rho \geqq 0$, then the function

$$\varphi(\rho) = \psi((\rho + \nu^2)^{\frac{1}{2}} - \nu) \tag{40}$$

is again completely monotone in $\rho \geqq 0$.

When dealing with the sphere (38) we shall restrict ourselves to the class of functions (40), the constant ν having the value

$$\nu = \frac{k-1}{2}.$$

This class is rather narrow; nevertheless it contains the functions (8) for $0 < c < \nu^2$. In fact, the corresponding function $\psi(\rho)$ is

$$[(\rho + \nu)^2 + c - \nu^2]^{-1}$$

and this is identical with

$$(\nu^2 - c)^{-\frac{1}{2}} \int_0^\infty e^{-\rho t} e^{-\nu t} \sinh (\nu^2 - c)^{\frac{1}{2}} t \, dt. \tag{41}$$

7. In our present case the Green's function $G(\xi, x)$ depends only on the geodesic distance between the two points ξ, x, both on the sphere (38). Denoting this distance by r, $0 < r \leqq \pi$, we have as an analogue to (10) the formula

$$G(r) = \sum_{n=0}^{\infty} (n + \nu)\varphi(n(n + 2\nu))P_n^{(\nu)}(\cos r). \tag{42}$$

The spherical harmonics $P_n^{(\nu)}$ can be defined either by

$$\frac{1}{\nu} \sum_{n=0}^{\infty} P_n^{(\nu)}(\cos r) w^n = (1 - 2w \cos r + w^2)^{-\nu} \tag{43}$$

or by

$$\sum_{n=0}^{\infty} (n + \nu) P_n^{(\nu)}(\cos r) w^n = (1 - w^2) \cdot (1 - 2w \cos r + w^2)^{-\nu-1}. \tag{44}$$

Now, by (40) and (39),

$$\varphi(n(n + 2\nu)) = \psi(n) = \int_0^\infty e^{-nt} \, d\alpha(t)$$

and therefore, by (44),

$$G(r) = \int_0^\infty \frac{(1 - e^{-2t}) \, d\alpha(t)}{(1 - 2 \cos r \, e^{-t} + e^{-2t})^{\nu+1}}. \tag{45}$$

Since

$$(46) \qquad 1 - 2\cos r \cdot e^{-t} + e^{-2t} = (1 - e^{-t})^2 + 2(1 - \cos r)e^{-t},$$

the function $G(r)$ is monotonely non-decreasing for decreasing values of r. About its behavior in the neighborhood of $r = 0$ we shall deduce the following theorem.

THEOREM. *The assertions* I, II, III *of* §3 *are true if in their wording the function* $H(r)$ *is replaced by the function* (42) *and the function* $\varphi(\rho)$ *by the function* (40).

We first observe that for the present theorem the function (45) is equivalent to the function

$$(47) \qquad T(r) = \int_0^\infty \frac{t\, d\alpha(t)}{(t^2 + r^2)^{\nu+1}}.$$

This follows easily from the fact that, for each $\epsilon > 0$, the differences between the functions $G(r)$ and $T(r)$ and the corresponding functions

$$G_\epsilon(r) = \int_0^\epsilon \frac{(1 - e^{-2t})\, d\alpha(t)}{(1 - 2\cos r\, e^{-t} + e^{-2t})^{\nu+1}}, \qquad T_\epsilon(r) = \int_0^\epsilon \frac{t\, d\alpha(t)}{(t^2 + r^2)^{\nu+1}}$$

are bounded in $0 < r \leqq \pi/2$, and that

$$1 - \eta(\epsilon) \leqq \varliminf_{r\to 0} \frac{G_\epsilon(r)}{2T_\epsilon(r)} \leqq \varlimsup_{r\to 0} \frac{G_\epsilon(r)}{2T_\epsilon(r)} \leqq 1 + \eta(\epsilon),$$

where $\lim_{\epsilon\to 0} \eta(\epsilon) = 0$, since (see (46))

$$\lim_{t\to 0} \frac{1 - e^{-t}}{t} = \lim_{t\to 0} \frac{1 - e^{-2t}}{2t} = \lim_{r\to 0} \frac{2(1 - \cos r)}{r^2} = 1.$$

Now, since $\nu + 1 = \frac{1}{2}(k + 1)$, the function (47) differs only by a numerical constant from

$$(48) \qquad \int_{-\infty}^{\infty} \cdots \int \psi[(n_1^2 + \cdots + n_k^2)^{\frac{1}{2}}] \exp[2\pi i(n_1 x_1 + \cdots + n_k x_k)]\, dn_1 \cdots dn_k,$$

where $x_1^2 + \cdots + x_k^2 = r^2$ and $\psi(\rho)$ is our present function (39).[9] But (48) is the function (18) with $\psi(\rho^{\frac{1}{2}})$ in the place of $\varphi(\rho)$. Hence the assertions I, II, III are true if we replace $H(r)$ by (42) and $\varphi(\rho)$ by $\psi(\rho^{\frac{1}{2}})$. Obviously, for $\rho \to \infty$, the function $\psi(\rho^{\frac{1}{2}})$ can be replaced by the function (40), and this completes the proof of our theorem.

In a similar fashion we could prove that $\varphi(\Delta)f$ transforms analytic functions into analytic functions and that the function $\psi(\rho)$ need not be bounded in the whole half-line $0 \leqq \rho < \infty$.

[9] See S. Bochner, *Fouriersche Integrale*, p. 189, formula (21).

8. For $\varphi(\rho) = (\rho + c)^{-1}$, $0 < c < \nu^2$ we obtain for (45) (see (41))

$$G(r) \approx \int_0^\infty \frac{\sinh t \cdot \sinh (\nu^2 - c)^{\frac{1}{2}} t \cdot dt}{(\cosh t - \cos r)^{\nu+1}},$$

where "$\approx$" denotes equality up to a factor independent of r (but dependent on k and c). Hence, by partial integration,

$$G(r) \approx \int_0^\infty \frac{\cosh (\nu^2 - c)^{\frac{1}{2}} t\, dt}{(\cosh t - \cos r)^{\nu}}. \tag{49}$$

This formula remains valid for $c > \nu^2$, in which case

$$G(r) \approx \int_0^\infty \frac{\cos (c - \nu^2)^{\frac{1}{2}} t\, dt}{(\cosh t - \cos r)^{\nu}}.$$

For $\nu = \frac{1}{2}$, that is, $k = 2$, we hence obtain, writing $c - \nu^2 = p^2$,

$$G(r) \approx Q_{-\frac{1}{2}+pi} (\cos r) + Q_{-\frac{1}{2}-pi} (\cos r),$$

where $Q_n(\cos r)$ is a spherical harmonic of the second kind.[10] For $0 < c < \frac{1}{4}$, the connection between $G(r)$ and the standard function $Q_n(\cos r)$ is more involved.

For k even and ≥ 4 we put in (49), $(\nu^2 - c)^{\frac{1}{2}} = n + \frac{1}{2}$, $u = \sinh \frac{1}{2}t$, $z = \sin \frac{1}{2}r$, and obtain

$$G(r) \approx \int_0^\infty \frac{\Phi_n(u)\, du}{(u^2 + z^2)^{\nu}},$$

where

$$\Phi_n(u) = \frac{\cosh (n + \frac{1}{2})t}{\cosh \frac{1}{2}t}.$$

If n is an integer, that is, for

$$c = \left(\frac{k-1}{2}\right)^2 - \left(\frac{2n+1}{2}\right)^2, \qquad n = 0, 1, \cdots, \frac{k-4}{2}, \tag{50}$$

$\Phi_n(u)$ has the special form

$$a_0 + a_1 u^2 + \cdots + a_n u^{2n},$$

and therefore

$$G(r) \approx \frac{b_0}{z^{k-2}} + \frac{b_1}{z^{k-4}} + \frac{b_2}{z^{k-6}} + \cdots + \frac{b_n}{z^{k-2-2n}};$$

in particular

$$G(r) \approx \frac{1}{r^{k-2}} (\alpha_0 + \alpha_1 r^2 + \alpha_2 r^4 + \alpha_3 r^6 + \cdots).$$

[10] E. W. Hobson, *The Theory of Spherical and Ellipsoidal Harmonics*, 1931, p. 275, formula (149).

This expansion is remarkable for the absence of terms containing $\log r$ which one would expect to be present in an expansion of $G(r)$ around the origin. It can be shown that for values of c other than (50) the logarithmic terms actually do occur, but a further investigation of this question would have but little to do with the topic of the present note.

PRINCETON UNIVERSITY.

AN ANALOGUE OF THE VON STAUDT-CLAUSEN THEOREM

By Leonard Carlitz

1. **Introduction.** Let $GF(p^n)$ denote a fixed Galois field, and x an indeterminate over the field. The function[1]

$$\psi = \psi(t) = \sum_{i=0}^{\infty} \frac{(-1)^i}{F_i} t^{p^{ni}}, \tag{1.1}$$

where

$$[i] = x^{p^{ni}} - x, \qquad F_i = [i][i-1]^{p^n} \cdots [1]^{p^{n(i-1)}}, \qquad F_0 = 1, \tag{1.2}$$

is closely connected with the arithmetic of polynomials in the $GF(p^n)$. In this paper we study the coefficients in the reciprocal of (1.1), more precisely in t/ψ. In particular we shall be interested in proving an analogue of the von Staudt-Clausen theorem for these coefficients.

In order to define properly the coefficients in the reciprocal it is necessary to define a "normalizing" factor (analogous to $n!$ in ordinary arithmetic). This is done in the following way. Let

$$m = \alpha_0 + \alpha_1 p^n + \cdots + \alpha_s p^{ns} \qquad (0 \leqq \alpha_i < p^n)$$

be the canonical expansion of m to the base p^n; then we put

$$g(m) = F_0^{\alpha_0} F_1^{\alpha_1} \cdots F_s^{\alpha_s}, \qquad g(0) = 1, \tag{1.3}$$

where F_i has the same significance as in (1.2). Thus for example

$$g(p^{ns}) = F_s, \qquad g(p^{ns} - 1) = (F_1 \cdots F_{s-1})^{p^n - 1}.$$

We may now define the coefficients of the reciprocal by means of

$$\frac{t}{\psi} = \sum_{m=0}^{\infty} \frac{B_m}{g(m)} t^m, \tag{1.4}$$

the summation obviously containing only terms in which m is a multiple of $p^n - 1$. Clearly $B_0 = 1$ and B_m is a rational function of x. The analogy between B_m and the ordinary Bernoulli numbers is brought out by the relation[2]

$$\sum \frac{1}{E^m} = \frac{B_m}{g(m)} \xi^m \qquad (p^n - 1 | m),$$

where the summation is over *all* primary polynomials E, and

$$\xi = \lim_{k=\infty} \frac{[1]^{p^{nk}/(p^n-1)}}{[k][k-1] \cdots [1]}.$$

Received July 21, 1937.

[1] See this Journal, vol. 1 (1935), pp. 137–168. This paper will be cited as DJ.

[2] DJ, p. 161, Theorem 9.3.

In the present paper we discuss some of the arithmetic properties of B_m. Our principal result is the analogue of the von Staudt-Clausen theorem for the Bernoulli numbers. We find that

$$B_m = G_m - e \sum_{\deg P = k} \frac{1}{P} \qquad (p^n \neq 2),$$

where G_m is some polynomial (whose precise form is not determined), e is an integer, not divisible by p, and the summation is over all irreducible polynomials P of degree k; finally k is a number depending on m whose existence depends on a certain set of conditions (see (7.3) below); if the conditions are not satisfied $B_m = G_m$ and is therefore a polynomial. When $p^n = 2$ the result must be modified slightly.

The method of proof depends on certain ideas due to A. Hurwitz.[3] While the proof is not particularly difficult, there are a number of details that make it rather long. In particular it is necessary to prove certain lemmas on $g(m)$ which are of some interest in themselves.

2. Lemmas on $g(m)$.

THEOREM 1. *For $m_1, m_2 \geqq 0$, the quotient*

$$\frac{g(m_1 + m_2)}{g(m_1)g(m_2)} \tag{2.1}$$

is integral (that is, a polynomial).[4]

Let

$$\begin{aligned} m_1 &= \beta_0 + \beta_1 p^n + \cdots + \beta_s p^{ns} \qquad (0 \leqq \beta_i < p^n), \\ m_2 &= \gamma_0 + \gamma_1 p^n + \cdots + \gamma_s p^{ns} \qquad (0 \leqq \gamma_i < p^n); \end{aligned} \tag{2.2}$$

then

$$m_1 + m_2 = (\beta_0 + \gamma_0) + \cdots + (\beta_s + \gamma_s)p^{ns} \qquad (0 \leqq \beta_i + \gamma_i < 2p^n). \tag{2.3}$$

If now we put

$$\beta_0 + \gamma_0 = \alpha_0 + \delta_0 p^n,$$

where $\delta_0 = 0$ or 1, and $0 \leqq \alpha_0 < p^n$, we may define $\delta_1, \cdots, \delta_s$ recursively by means of

$$\begin{aligned} \delta_0 + \beta_1 + \gamma_1 &= \alpha_1 + \delta_1 p^n, \\ \delta_1 + \beta_2 + \gamma_2 &= \alpha_2 + \delta_2 p^n, \\ &\cdots\cdots\cdots\cdots \\ \delta_{s-1} + \beta_s + \gamma_s &= \alpha_s + \delta_s p^n, \end{aligned} \tag{2.4}$$

[3] Mathematische Annalen, vol. 51 (1899), pp. 196–226 (= Mathematische Werke II, Basel, 1933, pp. 342–373).

[4] Throughout this paper the word "integral" will be used to denote a polynomial in x with coefficients in $GF(p^n)$.

where each $\delta_i = 0$ or 1, and $0 \leqq \alpha_i < p^n$. Thus (2.3) becomes

$$m_1 + m_2 = \alpha_0 + \alpha_1 p^n + \cdots + \alpha_s p^{ns} + \delta_s p^{n(s+1)}.$$

Hence by the definition of $g(m)$ in (1.3),

$$g(m_1 + m_2) = F_1^{\alpha_1} \cdots F_s^{\alpha_s} F_{s+1}^{\delta_s}.$$

Comparison with (2.2) leads at once to

$$\frac{g(m_1 + m_2)}{g(m_1)g(m_2)} = F_{s+1}^{\delta_s} \prod_{j=1}^{s} F_j^{\alpha_j - \beta_j - \gamma_j}.$$

Since by (1.2) $F_{s+1} = [s + 1]F_s^{p^n}$, the right member

$$= [s + 1]^{\delta_s} F_s^{\delta_s p^n + \alpha_s - \beta_s - \gamma_s} \prod_{j=1}^{s-1} F_j^{\alpha_j - \beta_j - \gamma_j}$$

$$= [s + 1]^{\delta_s} F_s^{\delta_{s-1}} \prod_{j=1}^{s-1} F_j^{\alpha_j - \beta_j - \gamma_s},$$

by the last of (2.4). Again

$$F_s^{\delta_{s-1}} F_{s-1}^{\alpha_{s-1} - \beta_{s-1} - \gamma_{s-1}} = [s]^{\delta_{s-1}} F_{s-1}^{\delta_{s-2}}.$$

Proceeding in this way we have finally

$$\frac{g(m_1 + m_2)}{g(m_1)g(m_2)} = [s + 1]^{\delta_s} [s]^{\delta_{s-1}} \cdots [1]^{\delta_0}. \tag{2.5}$$

Thus we have not only proved that (2.1) is integral but have derived the explicit formula (2.5).

As an immediate corollary it is evident that *for* $m_1, \cdots, m_k = 0$, *the quotient*

$$\frac{g(m_1 + \cdots + m_k)}{g(m_1) \cdots g(m_k)} \tag{2.6}$$

is integral.

If in (2.6) we take $m_1 = \cdots = m_k = m$, we see that $g(km)/g^k(m)$ is integral. For later purposes it will be necessary to know that this quotient is divisible by $g(k)$. But we may prove without difficulty the following slightly stronger result.

THEOREM 2. *For* $m \geqq 0$, *define* $\mu = \mu(m)$ *by means of*

$$\begin{aligned} m &= \alpha_0 + \alpha_1 p^n + \cdots + \alpha_s p^{ns} \qquad (0 \leqq \alpha_i < p^n), \\ \mu &= \alpha_0 + \alpha_1 + \cdots + \alpha_s . \end{aligned} \tag{2.7}$$

Thus for $m \geqq 1$, $\mu \geqq 1$. *Then the quotient*

$$\frac{g(km)}{g^k(m) g^{\mu}(k)}$$

is integral.[5]

[5] Compare Bachmann, *Niedere Zahlentheorie.*

It will be convenient to take first the special case $m = p^{ni}$, $k = p^{nj}$, so that $\mu = 1$. We shall require the

LEMMA. *For all $i, j \geqq 0$, the quotient*

$$(2.8)\qquad (i, j) = \frac{F_{i+j}}{F_i^{p^{nj}} F_j} = \frac{[i+j][i+j-1]^{p^n} \cdots [i+1]^{p^{n(j-1)}}}{F_j}$$

is integral.

This is easily proved by induction. From (2.8) follows

$$(i, j) = \frac{[i+j]F_{i+j-1}^{p^n}}{F_i^{p^{nj}} F_j} = \frac{([j] + [i]^{p^{nj}})F_{i+j-1}^{p^n}}{F_i^{p^n} F_j}$$

$$= (i, j-1)^{p^n} + \left(\frac{F_{i+j-1}}{F_{i-1}^{p^{nj}}}\right)^{p^n - 1} (i-1, j).$$

Since $(i, 0) = 1 = (j, 0)$ it is evident from this recursion formula that (i, j) is integral for arbitrary non-negative i, j.

Returning to the general theorem, we now suppose m arbitrary but to begin with again take $k = p^{nj}$. From (2.7) and the definition of $g(m)$ it follows easily that

$$\frac{g(p^n m)}{g^{p^n}(m)} = [s+1]^{\alpha_s}[s]^{\alpha_{s-1}} \cdots [1]^{\alpha_0}.$$

Replace m by $p^n m$ and this becomes

$$\frac{g(p^{2n} m)}{g^{p^n}(p^n m)} = [s+2]^{\alpha_s}[s+1]^{\alpha_{s-1}} \cdots [2]^{\alpha_0};$$

combining the last two equations we have

$$\frac{g(p^{2n} m)}{g^{p^n}(p^n m)} = ([s+2][s+1]^{p^n})^{\alpha_s} \cdots ([2][1]^{p^n})^{\alpha_0}.$$

Continuing in this way we see that

$$(2.9)\qquad \frac{g(p^{nj} m)}{g^{p^{nj}}(m)} = \prod_{i=0}^{s} \{[i+j] \cdots [i+1]^{p^{n(j-1)}}\}^{\alpha_i}.$$

If we compare this with (2.8), it is clear that the right member is a multiple of

$$F_j^{\alpha_0 + \cdots + \alpha_s} = F_j^{\mu},$$

so that the theorem is proved for the special case $k = p^{nj}$.

In the next place, for $\beta \geqq 1$, it follows that

$$(2.10)\qquad F_j^{\beta\mu} \,\Big|\, \frac{g^{\beta}(p^{nj} m)}{g^{\beta p^{nj}}(m)}.$$

But by Theorem 1, $g^{\beta}(p^{nj}m) \mid g(\beta p^{nj}m)$, so that (2.10) implies

$$F_j^{\beta\mu} \,\Big|\, \frac{g(\beta p^{nj}m)}{g^{\beta p^{nj}}(m)}; \tag{2.11}$$

thus the theorem is proved for $k = \beta p^{nj}$, where $0 < \beta < p^n$.

Now take

$$k = \alpha p^{ni} + \beta p^{nj} \qquad (0 < \alpha, \beta < p^n;\ i \neq j). \tag{2.12}$$

Then by (2.11) we have

$$F_i^{\alpha\mu} F_j^{\beta\mu} \,\Big|\, \frac{g(\alpha p^{ni}m)\, g(\beta p^{nj}m)}{g^{\alpha p^{ni}}(m) g^{\beta p^{nj}}(m)};$$

but by (2.12) and (1.3)

$$g(\alpha p^{ni} + \beta p^{nj}) = F_i^{\alpha} F_j^{\beta},$$

and by Theorem 1,

$$g(\alpha p^{ni}m) g(\beta p^{nj}m) \mid g\{(\alpha p^{ni} + \beta p^{nj})m\},$$

so that

$$g^{\mu}(k) \,\Big|\, \frac{g(km)}{g^{k}(m)}$$

for k as in (2.12). Proceeding in this way we see that Theorem 2 is true generally.

3. ***H*-series.**[6] If in the series

$$S = \sum_{m=0}^{\infty} \frac{A_m t^m}{g(m)}, \tag{3.1}$$

the coefficients A_m are integral, we shall call (3.1) an H-series. It follows at once from the definition that if S and S' are H-series, then $AS + A'S'$ is also an H-series, where A and A' are any two polynomials in x alone. As for the product of two H-series, if A_m, A'_m are the coefficients of S and S' respectively, and C_m denotes the general coefficient in SS', we evidently have

$$C_m = \sum_{i+j=m} \frac{g(m)}{g(i)g(j)} A_i A'_j.$$

By Theorem 1, the g-quotients on the right are integral so that C_m is integral. Therefore the product of two H-series is itself an H-series, and generally for the product of any number of series.

Consider next the reciprocal of S. In general this is not an H-series. If however $A_0 = 1$—or any non-zero element of $GF(p^n)$—then the reciprocal is also an H-series. Thus for $A_0 = 1$, put $S = 1 + S_1$; then

$$\frac{1}{S} = \frac{1}{1 + S_1} = 1 - S_1 + S_1^2 - \cdots,$$

[6] Compare Hurwitz, loc. cit.

and it is clear from the above that the right member is an H-series. Finally it follows that in this case S'/S is also an H-series.

Of special interest is the series

$$S_1 = \sum_{m=1}^{\infty} \frac{A_m t^m}{g(m)}, \tag{3.2}$$

which has no constant term. We have seen above that for the k-th power,

$$S_1^k = \sum_{m=k}^{\infty} \frac{C_m t^m}{g(m)},$$

C_m is integral. We shall now show that C_m is a multiple of $g(k)$, or what amounts to the same thing, we prove

Theorem 3. *If S_1 is an H-series without constant term, then $S_1^k/g(k)$ is also an H-series.*

Assume first $k = p^{ni}$. Then by (3.2),

$$S_1^{p^{ni}} = \sum_1^{\infty} \frac{A_m^{p^{ni}} t^{mp^{ni}}}{g(p^{ni} m)} \frac{g(p^{ni} m)}{g^{p^{ni}}(m)}.$$

But by Theorem 2, the second fraction on the right is divisible by $g(p^{ni}) = F_i$. If then we put

$$S_1^{p^{ni}} = \sum \frac{C_m t^m}{g(m)}, \tag{3.3}$$

we have $F_i \mid C_m$. Squaring both sides of (3.3), we write

$$S_1^{2p^{ni}} = \sum \frac{C'_m t^m}{g(m)}, \quad \text{where } C'_m = \sum_{e+f=m} \frac{g(m)}{g(e)g(f)} C_e C_f.$$

Thus it is clear that $F_i^2 \mid C'_m$, in other words $S_1^{2p^{ni}}/F_i^2$ is an H-series. Similarly for $S^{\alpha p^{ni}}/F_i^{\alpha}$. In other words, $S_1^k/g(k)$ is an H-series for $k = \alpha p^{ni}$, $0 < \alpha < p^n$.

Suppose next that

$$k = \alpha p^{ni} + \beta p^{nj}, \qquad 0 < \alpha, \beta < p^n, \; i \neq j. \tag{3.4}$$

Put

$$S_1^{\alpha p^{ni}} = \sum \frac{A'_m t^m}{g(m)}, \qquad S_1^{\beta p^{nj}} = \sum \frac{A''_m t^m}{g(m)},$$

so that

$$F_i^{\alpha} \mid A'_m, \qquad F_j^{\beta} \mid A''_m. \tag{3.5}$$

Then for

$$S_1^k = \sum \frac{C_m t^m}{g(m)}, \qquad C_m = \sum_{e+f=m} \frac{g(m)}{g(e)g(f)} A'_e A''_f,$$

so that by (3.5)

$$F_i^{\alpha} F_j^{\beta} \mid C_m.$$

But by (3.4)

$$F_i^{\alpha} F_j^{\beta} = g(\alpha p^{ni} + \beta p^{nj}) = g(k),$$

so that the theorem is proved in this case also. It is now clear how the theorem may be proved for general k.

As a corollary of some interest, we state: *if an H-series without constant term be substituted for t in the H-series* (3.1) *and the result written as a series in t, then this series is also an H-series.*

4. **Some theorems on B_m.** If we define $A_m^{(k)}$ by means of

$$\psi^{p^{nk}-1} = \sum_{m=p^{nk}-1}^{\infty} \frac{A_m^{(k)} t^m}{g(m)},$$

then we have the formula[7]

$$B_m = \sum_{p^{nk} \leq m+1} \frac{A_m^{(k)}}{L_k}, \tag{4.1}$$

where

$$L_k = [k][k-1] \cdots [1], \qquad L_0 = 1. \tag{4.2}$$

Now $A_m^{(k)}$ is integral and by Theorem 3 is a multiple of $g(p^{nk} - 1)$. But by (1.3) and (4.2),

$$\frac{g(p^{nk}-1)}{L_k} = \frac{(F_k \cdots F_1)^{p^n - 1}}{L_{k-1}} \frac{1}{[k]}. \tag{4.3}$$

If we recall that $[k]$ is the product of the irreducible polynomials whose degree divides k, it is clear that if the left member of (4.3) be reduced to its lowest terms, then the factors of the denominator are simple. Further, except for the case $p^n = 2 = k$, *the irreducible factors are all of degree k.* This proves the following

THEOREM 4. *If $B_m = N_m/D_m$, where N_m and D_m are relatively prime, then D_m has only simple factors.*

In the next place from the identity[8]

$$x\psi(t) - \psi(xt) = \psi^{p^n}(t) \tag{4.4}$$

follows

$$\frac{xt}{\psi(xt)} - \frac{t}{\psi(t)} = t\frac{x\psi(t) - \psi(xt)}{\psi(xt)\psi(t)} = \frac{t\psi^{p^n-1}(t)}{\psi(xt)} = \frac{t\psi^{p^n-2}(t)}{x - \psi^{p^n-1}(t)} = \sum \frac{C_m}{g(m)} \frac{t^m}{x^m}, \tag{4.5}$$

where, by the discussion of §3, the coefficients C_m are clearly integral. Hence by (1.4) and the last theorem, it follows that *the product $x(x^m - 1)B_m$ is integral.*

[7] DJ, p. 158, formula (8.11).
[8] DJ, p. 150, formula (5.09).

But this result may be extended considerably. If G is an arbitrary polynomial of degree s, say, then in place of (4.4) we have the more general formula[9]

$$(4.6) \qquad G\psi(t) - \psi(Gt) = \sum_{j=1}^{s} A_j \psi^{p^{nj}-1}(t),$$

where the A_j are integral; indeed

$$A_j = \frac{(-1)^{j-1}}{F_j} \psi_j(G), \qquad \psi_j(u) = \sum_{i=0}^{j} (-1)^{j-i} \frac{F_j}{F_i L_{j-i}^{p^{ni}}} u^{p^{ni}}.$$

Thus (4.5) becomes

$$\frac{Gt}{\psi(Gt)} - \frac{t}{\psi(t)} = t\frac{G\psi(t) - \psi(Gt)}{\psi(Gt)\psi(t)} = \frac{t \sum A_j \psi^{p^{nj}-2}}{G - \sum A_j \psi^{p^{nj}-1}}.$$

Now both numerator and denominator on the right are H-series, and it is easily seen that if we replace t by Gt the quotient becomes an H-series. Thus applying Theorem 4, we have

THEOREM 5. *For G an arbitrary polynomial, the product $G(G^m - 1)B_m$ is integral.*

Assume now in the notation of Theorem 4 that $P \mid D_m$, where P is an irreducible polynomial of degree k. Since in the theorem just proved, G is quite arbitrary we may take it equal to a primitive root (mod P). Now by the theorem,

$$G^m \equiv 1 \pmod{P},$$

and therefore because of the nature of G, m must be a multiple of $p^{nk} - 1$. This proves

THEOREM 6. *If P is an irreducible divisor of the denominator of B_m, then $p^{nk} - 1$ divides m.*

If we return to the definition of $A_m^{(k)}$ and make use of (4.1) and (4.3), it is now evident that for $p^{nk} - 1$ not a divisor of m, $A_m^{(k)}$ is a multiple of P. Therefore in determining the fractional part of B_m it is necessary to retain in the right member of (4.1) only those terms for which $p^{nk} - 1 \mid m$. We may now state the following

THEOREM 7. *For P irreducible of degree k, we have the congruence*[10]

$$(4.7) \qquad \psi^{p^{nk}-1} \equiv \sum_{p^{nk}-1 \mid m} \frac{A_m^{(k)} t^m}{g(m)} \pmod{P}$$

the summation extending over multiples of $p^{nk} - 1$ only. The formula (4.1) *reduces to*

$$(4.8) \qquad B_m = G_m + \sum_{p^{nk}-1 \mid m} \frac{A_m^{(k)}}{L_k},$$

[9] DJ, p. 151, formula (5.11).

[10] The statement $\sum \frac{A_m t^m}{g(m)} \equiv \sum \frac{A'_m t^m}{g(m)}$ (mod P) is short for the infinite system of congruences $A_m \equiv A'_m$ (mod P).

where[11] G_m *is integral, and the summation extends over such* k *for which* $p^{nk} - 1$ *divides* m.

5. A lemma on $\psi^{p^{nk}-1}$. We now prove the following theorem, which is the most important point in the proof of the main theorem concerning B_m.

THEOREM 8. *For* P *irreducible of degree* k,

$$\psi^{p^{nk}-1} \equiv \left(\sum \frac{(-1)^{ki}}{F_{ki}} t^{p^{nki}}\right)^{p^{nk}-1} \pmod{P}. \tag{5.1}$$

In other words, in forming the $(p^{nk} - 1)$-th power of $\psi(t) \pmod{P}$ we may ignore those terms in $t^{p^{ni}}$ for which i is not a multiple of k. To prove (5.1), we remark first that by Theorem 3,

$$\psi^{p^{nk}} \equiv 0 \pmod{P}.$$

Combining this with (4.7), we have the congruence

$$\sum_{p^{nk}-1 \mid m} \frac{A_m^{(k)} t^m}{g(m)} \sum \frac{(-1)^i}{F_i} t^{p^{ni}} \equiv 0 \pmod{P}.$$

Picking out the coefficient of t^m on the left, we get

$$\sum_{p^{ni} \leq m} (-1)^i \frac{g(m)}{F_i g(m - p^{ni})} A_{m-p^{ni}}^{(k)} \equiv 0. \tag{5.2}$$

We suppose hereafter that $m - 1$ is a multiple of $p^{nk} - 1$. But by Theorem 7,

$$A_{m-p^{ni}}^{(k)} \equiv 0 \quad \text{for} \quad p^{nk} - 1 \nmid m - p^{ni},$$

that is, for i not a multiple of k. Therefore (5.2) becomes

$$\sum_{p^{nki} \leq m} (-1)^{ki} \frac{g(m)}{F_{ki} g(m - p^{nki})} A_{m-p^{nki}}^{(k)} \equiv 0. \tag{5.3}$$

Now if $p^{nk} \nmid m$ it is easily verified that the quotient of $g(m)$ by $g(m - 1) \not\equiv 0 \pmod{P}$. Indeed if $p^{ns} \mid m$, $p^{n(s+1)} \nmid m$, the quotient $= L_s$, as defined in (4.2). Hence for $p^{nk} \nmid m$, (5.3) becomes

$$A_{m-1}^{(k)} \equiv -\frac{1}{L_s} \sum_{1 < p^{nki} \leq m} (-1)^{ki} \frac{g(m)}{F_{ki} g(m - p^{nki})} A_{m-p^{nki}}^{(k)}, \tag{5.4}$$

the summation extending over $i > 0$ only.

We next recall a result proved elsewhere:[12]

$$A_{p^{nk}m-1}^{(k)} = 0 \quad \text{for} \quad m > 1, \tag{5.5}$$

which we shall apply to (5.3) in order to show that $A_m^{(k)} \equiv 0 \pmod{P}$ for a

[11] G_m will generally denote a polynomial depending on the index m and not necessarily the same in all formulas in which it occurs.

[12] DJ, p. 158, Theorem 8.3.

certain set of values of m. First we make a slight change in notation. In (5.4) replace m by $m + 1$, so that

$$(5.6) \qquad p^{nk} - 1 \mid m, \qquad p^{nk} \nmid m + 1.$$

Assume now that (5.6) holds but in addition $p^{nk} \mid m + 2$. Then clearly (5.4) shows that $A_m^{(k)}$ is a sum of terms of the type

$$A^{(k)}_{m+1-p^{nki}} \qquad (i > 0);$$

but according to (5.5), this vanishes unless, for some i, $m + 1 - p^{nki} = p^{nk} - 1$, that is, unless

$$(5.7) \qquad m = p^{nki} + (p^{nk} - 2)\cdot 1.$$

In the same way if

$$p^{nk} \nmid m + 1, \qquad p^{nk} \nmid m + 2, \qquad p^{nk} \mid m + 3$$

(so that $p^{nk} > 2$), then two applications of (5.4) lead to a linear homogeneous expression for $A_m^{(k)}$ in terms of the type

$$A^{(k)}_{m+2-p^{nki}-p^{nkj}} \qquad (i, j > 0),$$

which vanishes unless $m + 2 - p^{nki} - p^{nkj} = p^{nk} - 1$, that is, unless

$$(5.8) \qquad m = p^{nki} + p^{nkj} + (p^{nk} - 3)\cdot 1 \qquad (p^{nk} > 2).$$

Now in both (5.7) and (5.8) m is expressed as a sum of $p^{nk} - 1$ terms p^{nki}. We shall now show generally that unless

$$(5.9) \qquad m = p^{nki_1} + \cdots + p^{nki_r}, \qquad r = p^{nk} - 1,$$

$A_m^{(k)} \equiv 0 \pmod{P}$. For suppose

$$(5.10) \qquad \begin{aligned} &p^{nk} \mid m + i && (i = 1, \cdots, t), \\ &p^{nk} \nmid m + t + 1 && (t < p^{nk}). \end{aligned}$$

Apply (5.3) t times and $A_m^{(k)}$ is exhibited as a sum of terms

$$A_w^{(k)}, \qquad w = m + t - p^{nki_1} - \cdots - p^{nki_t}.$$

Since by the second of (5.10) $p^{nk} \mid w + 1$, follows from (5.5) that $A_w^{(k)} \equiv 0$ unless $w = p^{nk} - 1$, that is, unless

$$m = p^{nki_1} + \cdots + p^{nki_t} + (p^{nk} - t - 1)\cdot 1,$$

which is precisely the condition (5.9).

It is now easy to establish the congruence (5.1). If we expand $\psi^{p^{nk}-1}$ directly, it is clear that a term involving t^m occurs when

$$(5.11) \qquad m = p^{nj_1} + \cdots + p^{nj_r}, \qquad r = p^{nk} - 1.$$

But by the result just proved the sum of such terms for fixed m will be $\equiv 0 \pmod{P}$ unless m is of the form (5.9). But from this it follows almost imme-

diately that each j in (5.11) is a multiple of k. Hence in forming $\psi^{p^{nk}}$ it is necessary to use only the terms in $t^{p^{nki}}$, but as remarked at the beginning of this section, this is equivalent to (5.1). We have therefore established Theorem 8.

As we shall see in the next section it is now possible to evaluate $A_m^{(k)}$. In proving the von Staudt-Clausen theorem for the ordinary Bernoulli numbers, it is possible to make use of certain explicit formulas. Their analogues are not available, and therefore some such method as used here seems necessary.

6. **Further lemmas on $\psi^{p^{nk}-1}$.** We now consider

$$\left(\sum_{i=0}^{\infty} \frac{(-1)^{ki}}{F_{ki}} t^{p^{nki}}\right)^{p^{nk}-1}. \tag{6.1}$$

Rewrite (5.7) in the form

$$\begin{aligned} m &= \alpha_0 + \alpha_1 p^{nk} + \cdots + \alpha_s p^{nks} \qquad (\alpha_i \geqq 0), \\ p^{nk} - 1 &= \alpha_0 + \alpha_1 + \cdots + \alpha_s . \end{aligned} \tag{6.2}$$

Then it is clear that (6.1) becomes

$$\sum_m \frac{(p^{nk}-1)!}{\alpha_0!\,\alpha_1! \cdots \alpha_s!} \frac{(-1)^{k(\alpha_1+2\alpha_2+\cdots+s\alpha_s)}}{F_k^{\alpha_1} F_{2k}^{\alpha_2} \cdots F_{sk}^{\alpha_s}} t^m,$$

the summation extending over all m for which (6.2) is solvable. Making use of (5.1) we see that

$$A_m^k \equiv (-1)^{k(\alpha_1+\cdots+s\alpha_s)} \frac{(p^{nk}-1)!}{\alpha_0! \cdots \alpha_s!} \frac{g(m)}{F_k^{\alpha_1} \cdots F_{sk}^{\alpha_s}}, \tag{6.3}$$

where again m is of the form (6.2); for other m, $A_m^{(k)} \equiv 0$. All the congruences are (mod P), where as above P is irreducible of degree k. To determine when the multinomial coefficient in the right member of (6.3) is different from zero (that is, not a multiple of p) we use a theorem of Dickson's:[13] *If the coefficients α_i in* (6.2) *are of the form*

$$\alpha_i = \sum_{j=0}^{nk-1} \alpha_{ij} p^j \qquad (0 \leqq \alpha_{ij} \leqq p-1), \tag{6.4}$$

then

$$(\alpha_0, \cdots, \alpha_s) = \frac{(p^{nk}-1)!}{\alpha_0! \cdots \alpha_s!}$$

is prime to p if and only if

$$\sum_{i=0}^{s} \alpha_{ij} = p - 1 \qquad (j = 0, \cdots, nk-1). \tag{6.5}$$

[13] Annals of Mathematics, (1), vol. 11 (1896–97), pp. 75–76; Quarterly Journal of Mathematics, vol. 33 (1902), pp. 378–384.

If (6.5) *is satisfied, then*

$$(\alpha_0, \cdots, \alpha_s) \equiv \frac{(-1)^{nk}}{\prod_{i,j} \alpha_{ij}!} \pmod{p}. \tag{6.6}$$

We may therefore assume that m satisfies both (6.2) and (6.5).

It is now easy to evaluate the g-quotient in (6.3). Put

$$m = \beta_0 + \beta_1 p^n + \beta_2 p^{2n} + \cdots \qquad (0 \leqq \beta_i < p^n). \tag{6.7}$$

Comparison with (6.2) gives

$$\alpha_i = \sum_{j=0}^{k-1} \beta_{ki+j} p^{nj} \qquad (i = 0, \cdots, s). \tag{6.8}$$

Comparison with (6.4) gives

$$\beta_{ki+j} = \sum_{e=0}^{n-1} \alpha_{i,nj+e} p^e. \tag{6.9}$$

Now by (6.7), $g(m) = F_1^{\beta_1} F_2^{\beta_2} \cdots$, from which follows by use of (6.8),

$$\frac{g(m)}{F_k^{\alpha_1} \cdots F_{ks}^{\alpha_s}} \equiv (F_1^{\beta_1} \cdots F_{k-1}^{\beta_{k-1}})(F_1^{\beta_{k+1}} \cdots F_{k-1}^{\beta_{2k-1}}) \cdots$$

$$\equiv F_1^{\beta_1 + \beta_{k+1} + \cdots} F_2^{\beta_2 + \beta_{k+2} + \cdots} \cdots .$$

But by (6.9) and (6.4)

$$\sum_i \beta_{ki+j} = \sum_{e=0}^{n-1} \sum_i \alpha_{i,nj+e} p^e = p^n - 1,$$

so that the g-quotient

$$\equiv (F_1 \cdots F_{k-1})^{p^n - 1} = (-1)^{k-1}, \tag{6.10}$$

by Wilson's Theorem and the fact[14] that F_k is the product of the primary polynomials of degree k. Therefore by (6.3) and (6.6) we conclude that *for m satisfying both* (6.2) *and* (6.5), *we have*

$$A_m^{(k)} \equiv (-1)^{k-1+k(\alpha_1+\cdots+s\alpha_s)} \frac{(-1)^{nk}}{\prod \alpha_{ij}!} \pmod{P}; \tag{6.11}$$

in all other cases, $A_m^{(k)} = 0$.

Let m be fixed; we shall now show that *at most one value of k can be found for which* (6.2) *and* (6.5) *are simultaneously satisfied.* For assume the relations

$$m = \gamma_0 + \gamma_1 p^{nl} + \cdots + \gamma_t p^{tnl},$$

$$p^{nl} - 1 = \gamma_0 + \gamma_1 + \cdots + \gamma_t,$$

[14] Bulletin of the American Mathematical Society, vol. 38 (1932), pp. 736–744; also DJ, p. 140.

(6.12)
$$\gamma_i = \sum_{j=0}^{nl-1} \gamma_{ij} p^j \qquad (0 \leqq \gamma_{ij} \leqq p - 1),$$

$$p - 1 = \sum_{i=0}^{t} \gamma_{ij}.$$

Then from (6.2) and (6.5) follows

$$\sum_{i=0}^{s} \sum_{j=0}^{nk-1} \alpha_{ij} = (p - 1)nk;$$

on the other hand from (6.12) follows

$$\sum_{i=0}^{t} \sum_{j=0}^{nl-1} \beta_{ij} = (p - 1)nl.$$

But clearly the α_{ij} and β_{ij} coincide (except for numbering), and therefore $k = l$, as asserted above.

We remark that there may be no value of k for which (6.2) and (6.5) hold. In this event B_m has no fractional part.

7. **The main theorem.** We return to (4.1). In view of the last result in §6, (4.1) becomes

(7.1)
$$B_m = G_m + \frac{1}{L_k} A_m^{(k)}, \qquad (p^n \neq 2)$$

provided k exists satisfying both (6.2) and (6.5); otherwise (7.1) is simply $B_m = G_m$, so that B_m is integral and nothing further need be said. Assuming then that a k exists, we make use of (4.3). If we exclude for the moment the case $p^n = 2$, it follows that the irreducible divisors of the denominator of B_m are all of degree k. Now it is easy to show (see the remark immediately following (4.3)) that

$$-\frac{1}{[k]} = \sum_{\deg P \mid k} \frac{P'}{P},$$

where the summation extends over all irreducible P of degree a divisor of k, and P' denotes the derivative of P. By (4.3) we have

(7.2)
$$\frac{A_m^{(k)}}{L_k} = G_m - \frac{A_m^{(k)}}{L_{k-1}} \sum_{\deg P = k} \frac{P'}{P},$$

the summation now extending over irreducibles of degree k only. But[15]

$$(-1)^{k-1} L_{k-1} \equiv P' \qquad (\mathrm{mod}\ P),$$

so that (7.2) becomes

$$\frac{A_m^{(k)}}{L_k} = G_m - (-1)^{k-1} A_m^{(k)} \sum \frac{1}{P}.$$

[15] DJ, p. 166, formula (11.10).

Finally, making use of (6.11) and substituting in (7.1), we have

THEOREM 9 ($p^n \neq 2$). *For given m, the system*

$$(7.3)\qquad \begin{aligned} m &= \sum_{i=0}^{s} \alpha_i p^{nki}, & p^{nk-1} &= \sum_{i=0}^{s} \alpha_i, \\ \alpha_i &= \sum_{j=0}^{nk-1} \alpha_{ij} p^j, & p-1 &= \sum_{i=0}^{s} \alpha_{ij}, \qquad \alpha_{ij} \geqq 0, \end{aligned}$$

is either (i) *inconsistent, or* (ii) *consistent for a single value of k, in which case the* α_i, α_{ij} *are uniquely determined. In case* (i) B_m *is integral; in case* (ii) *we have*

$$(7.4)\qquad B_m = G_m - \frac{(-1)^{nk+k(\alpha_1+\cdots+s\alpha_s)}}{\prod_{i,j} \alpha_{ij}!} \sum_{\deg P=k} \frac{1}{P}.$$

Here G_m is integral and the summation is over all irreducible P of degree k.

It remains to consider the excluded case $p^n = 2$. We may no longer conclude from (4.3) that the denominator of B_m contains only irreducibles of degree k; polynomials of the first degree may also occur. To decide when this happens we examine

$$\psi^3 = \left(t + \frac{t^2}{F_1} + \frac{t^4}{F_2} + \cdots\right)\left(t^2 + \frac{t^4}{F_1^2} + \frac{t^8}{F_2^2} + \cdots\right).$$

Clearly $A_m^{(2)}$ is different from zero only when m is of the form $2^\alpha + 2^\beta > 2$. Since in this case

$$g(m) = \begin{cases} F_\alpha F_\beta & \text{for } \alpha \neq \beta, \\ F_{\alpha+1} & \text{for } \alpha = \beta, \end{cases}$$

it is evident that for $\alpha \neq \beta$,

$$\begin{aligned} A_m^{(2)} &= g(m)\left\{\frac{1}{F_\alpha F_{\beta-1}^2} + \frac{1}{F_\beta F_{\alpha-1}^2}\right\} \\ &= [\beta] + [\alpha] = x^{2^\beta} + x^2, \end{aligned}$$

while for $\alpha = \beta$,

$$A_m^{(2)} = g(m)\frac{1}{F_\alpha F_{\alpha-1}^2} = [\alpha+1][\alpha].$$

Thus for $\alpha = \beta$, it is clear that $A_m^{(2)}$ is a multiple of L_2.

For $\alpha \neq \beta$, there are several possibilities. If $\alpha > \beta > 0$, then $A_m^{(2)}$ is divisible by $(x^2 + x)^2$; but if $\alpha > \beta = 0$, then $A_m^{(2)}$ is divisible by $x^2 + x$ only and the quotient $A_m^{(2)}/(x^2 + x)$ is congruent to 1 (mod $x^2 + x$). Again for $\alpha \equiv \beta \pmod 2$, $A_m^{(2)}$ is divisible by $x^2 + x + 1$, while for $\alpha \not\equiv \beta \pmod 2$, $A_m^{(2)} \equiv 1$ (mod $x^2 + x + 1$). We now note that if the system (7.3) is satisfied for $p^n = 2 = k$, then it follows at once that $m = 4^i + 2\cdot 4^j$; in other words, this is the case $\alpha \not\equiv \beta \pmod 2$. Also it is easily seen that any other value of k is inconsistent

with $m = 2^\alpha + 2^\beta$ $(\alpha > \beta)$. Hence we have the following supplement to Theorem 9:

THEOREM 10 $(p^n = 2)$. *If the system* (7.3) *is consistent for* $k \neq 2$, *then*

$$B_m = G_m + \sum_{\deg P = k} \frac{1}{P}; \tag{7.5}$$

if $k = 2$, *then for* m *even*,

$$B_m = G_m + \frac{1}{x^2 + x + 1}, \tag{7.6}$$

while for m *odd*,

$$B_m = G_m + \frac{1}{x^4 + x} = G_m + \frac{1}{x} + \frac{1}{x+1} + \frac{1}{x^2 + x + 1}. \tag{7.7}$$

If (7.3) *is consistent for no value of* k, *then for* m *even*, $B_m = G_m$, *while for* m *odd*,

$$B_m = G_m + \frac{1}{x} + \frac{1}{x+1}.$$

The following remark may be useful in testing (7.3). We assume m fixed; pick a k such that $p^{nk} - 1 \mid m$. Then (because of the condition $0 < \alpha_i < p^{nk}$) the α_i are uniquely determined. If their sum is not $p^{nk} - 1$, we go no further. If however the sum is equal to $p^{nk} - 1$, we use the equation $\alpha_i = \sum \alpha_{ij} p^j$ to determine the α_{ij} (because of the condition $0 < \alpha_{ij} < p$, the determination is unique). It is then only necessary to check the system of equations $p - 1 = \sum_i \alpha_{ij}$.

A partial check on Theorems 9 and 10 is furnished by the case $m = p^{nk} - 1$, for here a simple explicit formula[16] is available for B_m:

$$B_m = \frac{g(m)}{L_k} = \frac{(F_1 \cdots F_{k-1})^{p^n - 1}}{L_k}. \tag{7.8}$$

For $p^n = 2 = k$, this reduces to

$$B_3 = \frac{1}{[2]} = \frac{1}{x} + \frac{1}{x+1} + \frac{1}{x^2 + x + 1},$$

which agrees with (7.7). In all other cases the irreducible divisors of the denominator of B_m are of degree k, and it is evident that (7.8) is in agreement with (7.5). For this value of m, it is clear that $s = 0$, $\alpha_0 = p^{nk} - 1$, $\alpha_{ij} = p - 1$, $A_m^{(k)} = g(m)$.

DUKE UNIVERSITY.

[16] DJ, p. 159, formula (9.02).

SUMS OF VALUES OF A POLYNOMIAL MULTIPLIED BY CONSTANTS

By Kenneth S. Ghent

1. **Introduction.** We seek conditions on the integer s, on the sets of positive integers $(a_1, \cdots, a_s)$ and on the coefficients of a polynomial $P(x)$ for which the Diophantine equation

$$n = \sum_{\nu=1}^{s} a_\nu P(h_\nu) \tag{1}$$

is solvable in integers $h_\nu \geqq 0$ for every integer n sufficiently large. By n sufficiently large we mean that n is greater than an existing constant b_1 which depends only on s, $a_1, \cdots, a_s$ and on the degree k and the coefficients of the polynomial $P(x)$. We consider two cases of the polynomial $P(x)$:

$$P(x) = a(x^3 - x)/6 + b(x^2 - x)/2 + cx + d, \tag{2}$$

where $a > 0$ and a, b, c and d are integers;[1]

$$\begin{aligned} P(x) = {} & ax(x+1)(x+2)(x+3)/24 + bx(x+1)(x+2)/6 \\ & + cx(x+1)/2 + dx + e, \end{aligned} \tag{3}$$

where $a > 0$ and a, b, c, d and e are integers.[2]

For $a_1 = \cdots = a_s = 1$, the problem is the classical Waring problem for third and fourth degree polynomials. If $P(x)$ in (2) is such that $a \not\equiv 4c \pmod 8$ James has shown that every sufficiently large integer n is a sum of nine values of $P(x)$.[3] We show that for $s = 9$ and for integral constants $(a_1, \cdots, a_s)$ satisfying certain congruential conditions given later, every sufficiently large integer can be expressed in the form (1). For $P(x)$ as in (3), Miss Humphreys has given conditions which are sufficient to prove that every sufficiently large integer n is expressible as the sum of 21 values of $P(x)$. Under certain further assumptions on $P(x)$ and on the sets of positive integers $(a_1, \cdots, a_s)$ we shall

Received June 26, 1936; in revised form, June 28, 1937.

[1] A cubic polynomial in x is an integer for all integers $x \geqq 0$ if and only if it is of the form (2)—R. D. James, *The representation of integers as sums of values of cubic polynomials*, American Journal of Mathematics, vol. 56 (1934), pp. 303–315. See also D. H. Hilbert, *Über die Theorie der algebraischen Formen*, Mathematische Annalen, vol. 36 (1890), pp. 511–512.

[2] A fourth degree polynomial in x is an integer for all integers $x \geqq 0$ if and only if it is of the form (3). See M. G. Humphreys, *On the Waring problem with polynomial summands*, this Journal, vol. 1 (1935), pp. 361–375. The proof is accomplished by a slight modification of that of Hilbert in the paper previously cited.

[3] James, loc. cit.

show that every sufficiently large integer n can be expressed in the form (1) for $s = 21$.

The proof depends on certain analytic theorems and on certain congruential theorems. The analytic theorems follow closely those of James[4] and Landau.[5] We quote the analytic theorems without proof;[6] the congruential theory is given in detail.

2. **Notations.** The following notations are used throughout this paper. Further notations will be introduced when required.

$$K = 2^{k-1}; \qquad \mathfrak{K} = 1 - 1/K; \qquad f_\nu(x) = \sum_{h=0}^{\infty} x^{a_\nu P(h)};$$

$r(n) = r(n, k, s, a_1, \cdots, a_s; a, b, c, d, e) =$ the number of solutions of (1); ρ denotes a primitive q-th root of unity;

$$S(a_\nu, \rho) = \sum_{h=r+1}^{r+m} \rho^{a_\nu P(h)};$$

$$S_{\nu_\rho} = \sum_h \rho^{a_\nu P(h)},$$

where h ranges over a complete set of residues mod q;

$$A_0(q) = \frac{1}{q^s} \sum_{\rho(q)} \prod_{\nu=1}^{s} S_{\nu_\rho} \rho^{-n};$$

$\mathfrak{S}_0 = \mathfrak{S}_{0n} = \sum_{q=1}^{\infty} A_0(q)$ will be referred to as the Singular Series;

$$(a, b) = \text{g.c.d. of } a \text{ and } b.$$

3. **Principal analytic theorem.** Consider the following polynomial

$$\Phi(x) = \alpha_0 x^k + \alpha_1 x^{k-1} + \cdots + \alpha_k, \tag{4}$$

where $\alpha_0 > 0$ and $\alpha_0, \alpha_1, \cdots, \alpha_k$ are integers. The value of s considered will be that of the first Hardy-Littlewood theory. The notations of the previous section are used with $P(x)$ replaced by $\Phi(x)$. Consider also equation (1) with $P(x)$ replaced by $\Phi(x)$. The theorem which follows can then be proved for the polynomial $\Phi(x)$ of degree k by a generalization of the proof given by Landau.[7]

THEOREM 1. *For $s \geqq (k-2)K + 5$*

$$\left| r(n) - \frac{\alpha_0^{-s/k} \Gamma^s(1 + 1/k)}{(a_1 \cdots a_s)^{2/k} \Gamma(s/k)} \mathfrak{S}_0 n^{s/k-1} \right| < C_{20} n^{s/k-1-\mathfrak{K}_5},$$

[4] James, op. cit.

[5] E. Landau, *Über die neue Winogradoffsche Behandlung des Waringschen Problems*, Mathematische Zeitschrift, vol. 31 (1929), pp. 319–338.

[6] Proofs of the analytic theorems are given in the writer's doctoral dissertation at the University of Chicago, August, 1935.

[7] Landau, op. cit.

where C_{20} is a constant depending only on Φ, s and $\max a_\nu$ *and B_5 is a constant depending only on k.*

4. **Primitive solutions.** Consider $\Phi(x)$ as in (4). Let θ be the highest power of the prime p which divides every coefficient of $\Phi'(x)$. Then

$$\Phi_0(x) = p^{-\theta}\Phi'(x)$$

has at least one coefficient prime to p. We define γ as follows:

$$\begin{aligned}\gamma &= \theta + 1 \text{ for } p > 2,\\ &= \theta + 2 \text{ for } p = 2.\end{aligned}$$

Let $M_0(m) = M_0(m, n)$ denote the total number of solutions of

$$\sum_{\nu=1}^{s} a_\nu \Phi(x_\nu) \equiv n \pmod{m}, \qquad 0 \leqq x_\nu < n. \tag{5}$$

For $m = p^l$, let $N_0(p^l, n)$ denote the number of solutions of (5) for which $p \nmid$ g.c.d. of $\Phi_0(x_\nu)$. Let $N_{sp}(p^l, n)$ denote the number of solutions with $p \nmid$ g.c.d. $a_\nu\Phi_0(x_\nu)$. The last solutions defined will be referred to as primitive solutions. By generalizations of the work of Landau and James previously referred to, the following three lemmas can be proved.

Lemma 1. *If $l \geqq \gamma$ then*

$$N_{sp}(p^l) = p^{(l-\gamma)(s-1)}N_{sp}(p^\gamma).$$

Lemma 2.

$$M_0(m) = m^{s-1}\sum_{q\mid m} A_0(q).$$

Lemma 3. *If p_h denotes the h-th prime, then*

$$\sum_{q\mid p_1^l\cdots p_t^l} A_0(q) = \prod_{p\leqq p_t}\sum_{q\mid p^l} A_0(q).$$

The following lemma is also required.

Lemma 4. *For $\epsilon \geqq 0$*

$$A_0(q) > -E_2 q^{-b/4+\epsilon},$$

where E_2 is a constant depending at most on k, s, $\alpha_0, \cdots, \alpha_k$, $a_1, \cdots, a_s$.

5. **Application to the cubic and to the quartic.** The analytic theorem and the lemmas that we have stated apply to a polynomial $\Phi(x)$ of degree k with integral coefficients and leading coefficient positive. The polynomials $P(x)$ of (2) and (3) have leading coefficient positive but the coefficients are not integers. Hence, for these cases, let $Q(x) = Q(x; v, t) = P(xv + t)$, where $t \geqq 0$ and $v > 0$ are definite integers depending on the coefficients of $P(x)$ in the respective cases. We choose v so that it contains the least number of factors required to

make $Q(x)$ a polynomial with integral coefficients. The choice of t depends more specially on the case under consideration. For $Q(x)$ we define θ as before. For both the cubic and the quartic, $p > 3$ implies $\theta = 0$ since p does not divide all coefficients of $P(x)$.

6. **Introduction to the congruential theory.** The main portion of this paper is devoted to the development of certain congruential theorems[8] for the polynomials (2) and (3). We are then able, with the aid of the analytic theory introduced, to prove the principal results given in Theorems 3 and 5. We assume that the cubic is not of the type excepted by James[9] and that the quartic is not of the type excepted by Miss Humphreys.[10]

The following definition will be used: For a prime p and a positive integer l, a set of positive integers $(a_1, \cdots, a_s)$ has the property $S(p, p^l)$ if one of the set $(a_1, \cdots, a_s)$ which we may designate by a_j is prime to p and if the a_i $(i = 1, \cdots, s; i \neq j)$ are such that for every integer n there exists a $t < s$ for which

$$\sum_{i=1}^{t} a_i \equiv n \pmod{p^l}$$

is solvable.

7. **Results for cubic polynomials.** We seek conditions on the coefficients of $P(x)$ in (2) and on the constants $(a_1, \cdots, a_9)$ under which the equation (1) is solvable in integers $h_\nu \geqq 0$ for all sufficiently large integers n. We may evidently assume $d = 0$ in (2). We may also assume that a, b, c have no common factor other than unity; for if $p \mid a$, $p \mid b$ and $p \mid c$, then $p \mid P(x)$ and $\sum_{\nu=1}^{s} a_\nu P(x_\nu)$ would represent only multiples of p.

We first prove Theorem 2, a congruential theorem for the cubic; and then by Lemma 5 we obtain Theorem 3, the final theorem for the cubic. We state these theorems as follows for $P(x)$ as in (2).

THEOREM 2. *If the polynomial $P(x)$ and the set of positive integers $(a_1, \cdots, a_s)$ satisfy the following three conditions:*

(a) *for every prime $p > 3$, at least six of the positive integers $(a_1, \cdots, a_s)$ are prime to p;*

(b) *the set $(a_1, \cdots, a_s)$ has the properties $S(2, 2^3)$ and $S(3, 3^2)$;*

(c) $a \not\equiv 4c \pmod 8$, *i.e., the polynomial is not of the type excepted by James; then for every integer n and for $Q(x) = P(vx + t)$ there exist primitive solutions of the congruence*

$$\sum_{i=1}^{s \geqq 9} a_i Q(x_i) \equiv n \pmod{p^\gamma}. \tag{6}$$

[8] The methods of proof are similar to those of L. E. Dickson, *Cyclotomy, higher congruences and the Waring's problem*, American Jour. of Math., vol. 57 (1935), pp. 391–424.

[9] Loc. cit.

[10] Loc. cit.

THEOREM 3. *Let $s \geqq 9$ be an integer; and let $P(x)$ and the set of positive integers $(a_1, \cdots, a_s)$ be such that the conditions* (a), (b), *and* (c) *of Theorem 2 are satisfied. Then there exists a number C_{40} depending only on s, a, b, c and $a_1, \cdots, a_s$ such that every integer $n > C_{40}$ is expressible in the form*

$$n = \sum_{\nu=1}^{s \geqq 9} a_\nu P(x_\nu), \qquad x_\nu \geqq 0.$$

8. **Congruential theory for cubic polynomials.** We first consider (6) for primes $p > 3$. In §5 we saw that $\theta = 0$ for $p > 3$, hence $\gamma = 1$. Consider the polynomial $Q(x)$ in the form

$$Q(x) = A_1x^3 + A_2x^2 + A_3x, \tag{7}$$

where A_1, A_2, A_3 are integers whose g.c.d. is unity. Suppose, first, that $A_1 \equiv 0 \pmod p$. Then the congruence is a second degree congruence which can be treated directly. Hence suppose $A_1 \not\equiv 0 \pmod p$. Then we may evidently take $A_1 \equiv 1 \pmod p$ since, if we determine d by $dA_1 \equiv 1 \pmod p$, dn ranges with n in (6) over all residues $\pmod p$. Let $a_1, \cdots, a_6$ be six members of the set $(a_1, \cdots, a_s)$ which are prime to the given prime p. Then each a_i $(i = 1, \cdots, 6)$ satisfies one of the congruences in

$$a_i \equiv x^3, \qquad a_i \equiv gx^3, \qquad a_i \equiv g^2x^3 \pmod p, \tag{8}$$

where g is a fixed primitive root of p. Hence without loss of generality assume

$$a_1 \equiv a_2u_1^3, \qquad a_3 \equiv a_4u_3^3 \pmod p. \tag{9}$$

Case 1. $A_2A_3(a_1 + a_2)(a_3 + a_4) \not\equiv 0 \pmod p$. Let $x_2 = -u_1x_1$. Then

$$\begin{aligned} a_1Q(x_1) + a_2Q(x_2) &\equiv a_2(u_1^3x_1^3 + x_2^3) + A_2(a_2u_1^3x_1^2 + a_2x_2^2) + A_3(a_2u_1^3x_1 + a_2x_2) \\ &\equiv a_2(u_1^3x_1^3 - u_1^3x_1^3) + A_2a_2(u_1^3x_1^2 + u_1^2x_1^2) + A_3a_2(u_1^3x_1 - u_1x_1) \\ &\equiv A_2a_2(u_1^3 + u_1^2)x_1^2 + A_3a_2(u_1^3 - u_1)x_1 \pmod p. \end{aligned}$$

Similarly, let $x_4 = -u_3x_3$. Then

$$a_3Q(x_3) + a_4Q(x_4) \equiv A_2a_4(u_3^2 + u_3^3)x_3^2 + A_3a_4(u_3^3 - u_3)x_3 \pmod p.$$

Next[11] let $x_1 = X_1 + z$, where z is such that

$$(2A_2a_2(u_1^3 + u_1^2))z + A_3a_2(u_1^3 - u_1) \equiv 0 \pmod p.$$

This has one solution, since, from hypothesis, $2A_2a_2(u_1^3 + u_1^2)$ is not divisible by p. Similarly, set $x_3 = X_3 + v$, where v is such that

$$2A_2a_4(u_3^3 + u_3^2)v + A_3a_4(u_3^3 - u_3) \equiv 0 \pmod p.$$

[11] If $A_3a_2(u_1^3 - u_1) \equiv 0 \pmod p$, this transformation is unnecessary.

We have then essentially

$$\sum_{i=1}^{4} a_i Q(x_i) \equiv rx^2 + sy^2 \equiv n \pmod{p}, \tag{10}$$

where neither r nor s is divisible by p. This congruence is solvable for every integer n.

Case 2.[12] $A_2A_3 \not\equiv 0$, $a_1 + a_2 \equiv 0 \pmod{p}$. Let $x_1 = x_2 + l$, where l is so chosen that $3l + A_2 \not\equiv 0 \pmod{p}$. Then

$$\begin{aligned} a_1Q(x_1) + a_2Q(x_2) &= a_1Q(x_2 + l) + a_2Q(x_2) \\ &= (a_1 + a_2)Q(x_2) + a_1(x_2^2(3l + A_2) + x_2(3l^2 + 2lA_2 + A_3) \\ &\qquad + (l^3 + l^2A_2 + lA_3)); \end{aligned}$$

$a_1Q(x_1) + a_2Q(x_2) \equiv a_1(3l + A_2)x_2^2 + a_1(3l^2 + 2lA_2 + A_3)x_2 + \text{const} \pmod{p}$.
The remainder of the proof is as in Case 1.

Case 3. $A_2A_3 \equiv 0 \pmod{p}$. We show that in this case we can make a transformation, $x = X + l$, which carries $Q(x)$ to

$$\begin{aligned} Q_1(x) &= X^3 + A_2'X^2 + A_3'X + \text{const} \\ &= X^3 + X^2(3l + A_2) + X(3l^2 + 2A_2l + A_3) + \text{const} \end{aligned}$$

in which $A_2'A_3' \not\equiv 0 \pmod{p}$. Suppose first that A_2 and A_3 are both divisible by p. Choose l prime to p and we have $A_2'A_3'$ not divisible by p as desired. Suppose next that A_2 is divisible by p but that A_3 is not divisible by p. Now

$$(3l^2 + A_3) \equiv 0 \pmod{p} \tag{11}$$

has at most two incongruent solutions for l. Hence there are at least $(p - 2)$ incongruent values of l for which (11) is not satisfied. Any one of these $(p - 2)$ values of l makes $A_2'A_3' \not\equiv 0 \pmod{p}$. Finally, suppose that A_2 is not divisible by p but that A_3 is divisible by p. There is a unique value of $l \pmod{p}$ for which $3l + A_2$ is divisible by p. One of the remaining $(p - 1)$ incongruent residues $\pmod{p}$ satisfies

$$3l + 2A_2 \equiv 0 \pmod{p}.$$

Hence there exist $(p - 2)$ incongruent values of l for which

$$(3l + A_2)(3l + 2A_2) \not\equiv 0 \pmod{p}.$$

We use one of these values of l in the transformation and we obtain $A_2'A_3' \not\equiv 0 \pmod{p}$.

Hence in all cases we can satisfy the condition that A_2A_3 be not divisible by p.

We require a primitive solution of (6); i.e., at least one $a_sQ_0(x)$ prime to p. Suppose that no one of the $a_iQ(x_i)$ which we have used satisfies the desired condition. Then choose x_5 so that $a_5Q_0(x_5)$ is prime to p. This we can do

[12] If $a_3 + a_4 \equiv 0 \pmod{p}$, the treatment is the same as when $a_1 + a_2 \equiv 0 \pmod{p}$.

since we have just seen that if A_3 is divisible by p we can make a transformation on x to give $Q_1(x)$ in which A_3' is not divisible by p; and let x_5 be divisible by p. Then $a_5Q(x_5)$ is divisible by p while $a_5Q_0(x_5) \equiv a_5A_3 \pmod{p}$ is not divisible by p.

We have then shown that for $p > 3$ there exist primitive solutions of the congruence

$$\sum_{i=1}^{6} a_i Q(x_i) \equiv n \pmod{p}$$

if for every prime $p > 3$, $a_1, \cdots, a_6$ are all prime to p. We have thus proved Theorem 2 for primes $p > 3$.

For $p \leqq 3$, James used in the congruence (6) fewer than 9 equal values of the polynomial and, if necessary, one additional value of the polynomial which insured that the primitivity condition be satisfied in order to solve the congruence for an arbitrary integer n. If we assume that our set $(a_1, \cdots, a_s)$ has the properties $S(2, 2^3)$ and $S(3, 3^2)$, the discussion given by James applies to the present problem.[13] We have thus proved Theorem 3.

9. **Results for the quartic polynomial.** We consider $P(x)$ as in (3). As in the case of the cubic polynomial, we may assume that a, b, c, d and e have no common factor other than unity. We prove Theorem 4, a congruential theorem for the quartic similar to Theorem 2 for the cubic, and give in Theorem 5 our final results for the quartic.

THEOREM 4. *If the polynomial $P(x)$ and the set of positive integers $(a_1, \cdots, a_s)$ satisfy the following conditions:*

(a) *for every prime $p > 3$, eleven of the constants $(a_1, \cdots, a_s)$ are prime to p;*

(b) *the set $(a_1, \cdots, a_s)$ has the properties $S(2, 2^4)$ and $S(3, 3^2)$;*

(c) *the coefficient of x in the normal form*[14] *of the polynomial is not divisible by a prime of the form $4m + 3$ $(p > 3)$;*

(d) *the polynomial is not of the kind excepted by Miss Humphreys,*

then for every integer n and for $Q(x) = P(vx + t)$ there exist primitive solutions of the congruence

$$\sum_{i=1}^{s=21} a_i Q(x_i) \equiv n \pmod{p^7}. \tag{12}$$

THEOREM 5. *Let $s \geqq 21$ be an integer; and let $P(x)$, the quartic polynomial in* (3), *and the set of positive integers $(a_1, \cdots, a_s)$ be such that the conditions* (a), (b), (c), (d) *of Theorem 4 are satisfied. Then there exists a constant C_{41} depending on s, a, b, c, d and* max a_i *such that every integer $n > C_{41}$ is expressible in the form*

$$n = \sum_{\nu=1}^{s} a_\nu P(x_\nu).$$

[13] Our hypothesis makes Lemmas 12, 13, and 14 of James' paper available for our discussion. The use of Lemma 8 can be avoided. Sections 5 and 6 of James' paper then give the desired result.

[14] The normal form of the polynomial is discussed in the next section.

10. **Congruential theory for the quartic.** Suppose that $Q(x)$ is of the form

$$Q(x) = \beta_0 x^4 + \beta_1 x^3 + \beta_2 x^2 + \beta_3 x + \beta_4,$$

where β_i $(i = 0, 1, \cdots, 4)$ are integers and $\beta_0 > 0$. This we can assume after transformation of $P(x)$ in (3) by replacing x by $vx + t$. As in the case of the cubic we may assume that $\beta_0 \equiv 1 \pmod p$. Moreover, if k is prime to p (this is the case for $p > 3$) we replace x by $X + z$ where z is so chosen that the coefficient of X^3 is divisible by p. Finally, after this transformation, we may evidently assume that the constant term is zero and consider (12) for $Q(x)$ replaced by

$$Q_1(x) = x^4 + A_2 x^2 + A_3 x.$$

We call $Q_1(x)$ the normalized form of the polynomial $Q(x)$. Then condition (c) of Theorem 4 implies $A_3 \not\equiv 0$ (mod $p = 4m + 3$, $p > 3$).

We consider the congruence (12) for primes $p > 3$. Then $\theta = 0$ and $\gamma = 1$. Hence we consider

$$\sum_{i=1}^{s=21} a_i Q_1(x_i) \equiv n \pmod p, \tag{13}$$

where we assume that $(a_1, \cdots, a_{11})$ are prime to p. By Lemma 39 of a paper by Huston,[15] there exist primitive solutions of

$$\sum_{i=1}^{5} a_i h_i^4 \equiv n \pmod p.$$

Put $x_i = h_i y_1$ $(i = 1, \cdots, 5)$, $x_i = h_i y_2$ $(i = 6, \cdots, 10)$, where y_1 and y_2 are prime to p and where the h_i are so chosen that

$$\sum_{i=1}^{5} a_i h_i^4 \equiv 0 \equiv \sum_{i=6}^{10} a_i h_i^4 \pmod p \tag{14}$$

and in each congruence at least one $a_i h_i$ is prime to p.

Case 1. $A_2 A_3 \not\equiv 0 \pmod p$. Suppose first that the values of h_i $(i = 1, \cdots, 5)$ which we have chosen to satisfy $\sum_{i=1}^{5} a_i h_i^4 \equiv 0 \pmod p$ also satisfy $\sum_{i=1}^{5} a_i h_i^2 \equiv 0 \pmod p$. Then by choice of h_1 or $-h_1$ we can have $\sum_{i=1}^{5} a_i h_i$ not divisible by p. This is true since we may take $a_1 h_1$ prime to p and since if $(h_1, \cdots, h_5)$ satisfies

$$\sum_{i=1}^{5} a_i h_i^4 \equiv \sum_{i=1}^{5} a_i h_i^2 \equiv \sum_{i=1}^{5} a_i h_i \equiv 0 \pmod p,$$

then $(-h_1, h_2, \cdots, h_5)$ satisfies the first two congruences but insures that $\sum_{i=1}^{5} a_i h_i$ is not divisible by p. Hence

[15] R. E. Huston, *Asymptotic generalizations of Waring's theorem*, Proceedings of the London Mathematical Society, (2), vol. 39 (1935), pp. 82–115.

$$\sum_{i=1}^{5} a_i Q_1(x_i) \equiv \sum_{i=1}^{5} a_i h_i y_1 \equiv n \pmod{p},$$

which is solvable for every integer n. To satisfy the condition that the solution be primitive it may be necessary to add one more summand, $a_6Q_1(x_6)$, where x_6 is divisible by p while

$$a_6 Q_1'(x_6) \equiv a_6(4x^3 + 2A_2x + A_3) \not\equiv 0 \pmod{p},$$

since p does not divide A_3.

Suppose next that the values of h_i ($i = 1, \cdots, 10$) which we have chosen to satisfy (14) are such that

$$\sum_{i=1}^{5} a_i h_i^2 \not\equiv 0, \qquad \sum_{i=6}^{10} a_i h_i^2 \not\equiv 0 \pmod{p}.$$

Then

$$\begin{aligned} \sum_{i=1}^{10} a_i Q_1(x_i) \equiv A_2\Big(\sum_{i=1}^{5} a_i h_i^2\Big) y_1^2 + A_3\Big(\sum_{i=1}^{5} a_i h_i\Big) y_1 \\ + A_2\Big(\sum_{i=6}^{10} a_i h_i^2\Big) y_2^2 + A_3\Big(\sum_{i=6}^{10} a_i h_i\Big) y_2 \pmod{p}. \end{aligned} \tag{15}$$

Put $y_2 = Y_2 + Z_2$, where Z_2 is so chosen that the coefficient of Y_2 in (15) is divisible by p. Put $y_1 = Y_1 + Z_1$, where Z_1 is so chosen that the coefficient of Y_1 in (15) is divisible by p. We then have essentially

$$\sum_{i=1}^{10} a_i Q_1(x_i) \equiv rY_1^2 + sY_2^2 \equiv n \pmod{p},$$

which is solvable for every integer n since

$$r = A_2 \sum_{i=1}^{5} a_i h_i^2 \quad \text{and} \quad s = A_2 \sum_{i=6}^{10} a_i h_i^2$$

are both not divisible by $p > 3$. As before we may add one more summand, $a_{11}Q_1(x_{11})$, divisible by p to insure that the primitivity condition is satisfied.

Case 2. $A_2 \equiv 0$, $A_3 \not\equiv 0 \pmod{p}$. Evidently this case is equivalent to the first part of the preceding case since A_2 divisible by p has the same effect in the congruence as

$$\sum_{i=1}^{5} a_i h_i^2 \equiv 0 \pmod{p}.$$

Case 3. $A_2 \equiv A_3 \equiv 0 \pmod{p}$. Then

$$\sum_{i=1}^{5} a_i Q_1(x_i) \equiv a_1x_1^4 + a_2x_2^4 + \cdots + a_5x_5^4 \equiv n \pmod{p}$$

has a primitive solution by Lemma 29 of Huston's paper.

Case 4. $A_2 \not\equiv 0$, $A_3 \equiv 0 \pmod{p = 4m + 1}$. Choose h_i so that

$\sum_{i=1}^{5} a_i h_i^4 \equiv 0 \pmod p$ and $\sum_{i=1}^{5} a_i h_i^2 \not\equiv 0 \pmod p$. This can be done since $\xi^4 \equiv 1 \pmod p$ has at least one root ξ such that $\xi^2 \not\equiv 1 \pmod p$. Hence if $(h_1, h_2, \cdots, h_5)$ is a solution of

$$\sum_{i=1}^{5} a_i h_i^4 \equiv \sum_{i=1}^{5} a_i h_i^2 \equiv 0 \qquad \pmod p,$$

then $(\xi h_1, h_2, \cdots, h_5)$ is a solution of $\sum_{i=1}^{5} a_i h_i^4 \equiv 0 \pmod p$ but does not satisfy $\sum_{i=1}^{5} a_i h_i^2 \equiv 0 \pmod p$. Use this latter set of h_i. Then we can obtain

$$\sum_{i=1}^{5} a_i Q_1(x_i) + \sum_{i=6}^{10} a_i Q_1(x_i) \equiv A_2\left(\sum_{i=1}^{5} a_i h_i^2\right) y_1^2 + A_2\left(\sum_{i=6}^{10} a_i h_i^2\right) y_2^2 \equiv n \quad \pmod p,$$

which is solvable for every integer n since neither $r = A_2\left(\sum_{i=1}^{5} a_i h_i^2\right)$ nor $y = A_2\left(\sum_{i=6}^{10} a_i h_i^2\right)$ is divisible by p.

The proof fails when A_3 is divisible by a prime of the form $4m + 3$. This gives rise to the exception in the theorem. An additional summand, $a_{11}Q(x_{11})$, can be added if necessary to satisfy the condition that our solution be primitive.

Hence for $p > 3$ we have the congruence (12) solvable for $s \geqq 11$, with the one exception noted.

For $p \leqq 3$ we assume that our set has the properties $S(2, 2^4)$ and $S(3, 3^2)$. This assumption reduces the discussion of this case to that of Miss Humphreys. Her proof for the quartic case, $p \leqq 3$, $a_1 = a_2 = \cdots = a_s = 1$ then applies to the case under discussion. Hence we have Theorem 4.

11. **Proofs of Theorems 3 and 5**. If $\mathfrak{S}_{0n} \geqq \eta > 0$, where η is independent of n, it follows from Theorem 1 that $r(n) > 0$ when

$$n > C_{42} = ((C_{25}\alpha_0^{s/k}(a_1 \cdots a_s)^{2/k}\Gamma(s/k)/\eta\Gamma^s(1 + 1/k))^{1/B_5}.$$

The proofs of Theorems 3 and 5 are thus reduced to the proof of

LEMMA 5. *If for the cubic polynomial* $s \geqq 9$ *and if for the quartic polynomial* $s \geqq 21$, *then*

$$\mathfrak{S}_{0n} \geqq \eta > 0,$$

where η *is independent of* n.

Since $M_0(p^l) \geqq N_{sp}(p^l)$ we have from Lemmas 2 and 1 and Theorems 2 and 4 in the respective cases

$$\begin{aligned} \sum_{q \mid p^l} A_0(q) &= p^{-l(s-1)} M_0(p^l) \geqq p^{-l(s-1)} N_{sp}(p^l) \\ &= p^{-l(s-1)} p^{(l-\gamma)(s-1)} N_{sp}(p^\gamma) \geqq p^{-\gamma(s-1)} \geqq p^{-t(s-1)}. \end{aligned} \tag{16}$$

where $t = 4$ for $P(x)$ in (3) and $t = 3$ for $P(x)$ as in (2). The final inequality holds since in the respective cases $(l \geqq 4 \geqq \gamma)$, $(l \geqq 3 \geqq \gamma)$. Then by Lemma 4 with $\epsilon = 1/8$

$$(17) \quad \sum_{q \mid p^l} A_0(q) = 1 + \sum_{\lambda=1}^{\infty} A_0(p^\lambda) > 1 - E_2 \sum_{\lambda=1}^{\infty} p^{-9\lambda/8} = 1 - E_2(p^{9/8} - 1)^{-1}.$$

Finally, the singular series is convergent for $s \geqq (k-2)2^{k-1} + 5$ and hence by (16), (17) and Lemma 3

$$\begin{aligned}\mathfrak{S}_{0n} &= \lim_{l\to\infty} \sum_{q \mid p_1^l \cdots p_l^l} A_0(q) = \lim_{l\to\infty} \prod_{p \leqq p_l} \sum_{q \mid p^l} A_0(q) \\ &\geqq \prod_p \max\left(p^{-t(s-1)}, 1 - E_2(p^{9/8} - 1)^{-1}\right) \\ &\geqq \eta > 0.\end{aligned}$$

UNIVERSITY OF CHICAGO.

TRIGONOMETRIC APPROXIMATION IN THE MEAN

By E. S. Quade

The following theorem is stated without proof by G. H. Hardy and J. E. Littlewood.[1]

THEOREM. *The class* Lip(α, p) *is identical with the class of functions* $f(x)$ *approximable in the mean p-th power, with error* $O(n^{-\alpha})$, *by trigonometrical polynomials of degree* n.

They remark in addition: *This approximation may be made in general by the Fourier polynomials of* $f(x)$; *the case* $p = \infty$, *in which this is not true, is exceptional.*

The initial purpose of this paper is to examine the range of values of p and α for which this theorem and remark are true and to supply proofs. In doing this, related theorems are obtained in which the approximations are in terms of the metric of a more extensive space than L_p and in which the functions that measure the degree of approximation are more general than $n^{-\alpha}$. These theorems and their proofs parallel to a large extent the theorems given by de la Vallée Poussin[2] and Dunham Jackson[3] for the class Lip(α).

We assume throughout that our functions $f(x)$ are periodic with the period 2π. The functions $\Phi(u)$ and $\Psi(u)$ are of Young's type.[4] That is, $\Phi(u)$ is non-negative, convex, and satisfies the relations $\Phi(0) = 0$ and $\Phi(u)/u \to \infty$ as $u \to \infty$; $\Psi(u)$ has similar properties and is such that Young's inequality

$$uv \leqq \Phi(u) + \Psi(u), \qquad u, v \geqq 0$$

holds. Throughout the paper we will write $\Phi \mid u \mid$, $\Psi \mid u \mid$, for $\Phi(\mid u \mid)$, $\Psi(\mid u \mid)$.

If $f(x)$ is measurable and such that $\int_0^{2\pi} \Phi \mid f \mid dx$ exists, $f(x)$ is said to belong to the space $L_\Phi(0, 2\pi)$. If $f(x)$ is such that the product $f(x)g(x)$ is integrable for every $g(x) \in L_\Psi$, then $f(x) \in L_\Phi^*$. For this space

$$\|f\|_\Phi = \sup_g \left| \int_0^{2\pi} f(x)g(x)\, dx \right|$$

for all measurable $g(x)$ with $\rho_g \equiv \int_0^{2\pi} \Psi \mid g \mid dx \leqq 1$. This space[5] is linear,

Received October 5, 1936; in revised form, July 9, 1937.

[1] *A convergence criterion for Fourier series*, Math. Zeit., vol. 28 (1928), pp. 612-634; in particular, p. 633.

[2] *Leçons sur l'Approximation des Fonctions*, Paris, 1919. We shall refer to this treatise as (P).

[3] *The Theory of Approximation*, New York, 1930. We shall refer to this treatise as (D).

[4] A. Zygmund, *Trigonometrical Series*, Warsaw, 1935, §§4.11, 4.142. We shall refer to this treatise as (Z). It contains extensive bibliographical references to original sources.

[5] (Z), §4.541.

metric, and complete. If $f(x) \epsilon L_\Phi^*$, we put, for $\delta > 0$,

$$\omega_\Phi(\delta; f) = \sup_{0<|h|\leq\delta} \| f(x+h) - f(x) \|_\Phi .$$

For $p > 1$, L_p is a class L_Φ^*. In L_p, $p \geqq 1$,

$$\omega_p(\delta; f) = \sup_{0<|h|\leq\delta} \| f(x+h) - f(x) \|_p = \sup_{0<|h|\leq\delta} \left(\int_0^{2\pi} | f(x+h) - f(x) |^p \, dx \right)^{1/p}.$$

If $\omega_p(\delta; f) = O(\delta^\alpha)$, $\delta \to 0$, $f(x)$ is said to belong to the class[6] Lip(α, p). The limiting case of Lip(α, p), denoted Lip(α, ∞) is identical with Lip(α). For brevity, we shall write $\| f \|$ for $\| f \|_\Phi$ whenever it will not lead to confusion.

In order to prove the next theorem we need the following two lemmas.

Lemma A. *If $t_n(x)$ is a trigonometrical polynomial of degree n at most, then*

$$\| t_n' \|_\Phi \leqq 2n \| t_n \|_\Phi .$$

This lemma is a trivial extension of a theorem of Zygmund[7] that, when $\Phi(u)$ is a Young's function and $t_n(x)$ a trigonometrical polynomial of degree n, then

$$\int_0^{2\pi} \Phi \left| \frac{t_n'}{n} \right| dx \leqq \int_0^{2\pi} \Phi | t_n | \, dx.$$

If $t_n(x) \equiv 0$, $\| t_n' \| = \| t_n \| = 0$. Assuming $t_n(x) \not\equiv 0$, we have, since $\| t_n \| \geqq \int_0^{2\pi} | t_n | \, dx > 0$, by Young's inequality and the above theorem of Zygmund

$$\begin{aligned} \frac{\| t_n' \|}{\| t_n \|} &= n \sup_g \left| \int_0^{2\pi} \frac{t_n'(x)}{n \| t_n \|} g(x) \, dx \right| \\ &\leqq n \left[\int_0^{2\pi} \Phi \left| \frac{t_n'}{n \| t_n \|} \right| dx + \rho_g \right] \\ &\leqq n \left[\int_0^{2\pi} \Phi \left| \frac{t_n}{\| t_n \|} \right| dx + \rho_g \right]. \end{aligned}$$

Since $\rho_g \leqq 1$ and[8]

$$\int_0^{2\pi} \Phi \left| \frac{t_n}{\| t_n \|} \right| dx \leqq 1,$$

the lemma follows.

This lemma holds also for the cases[9] L_p, $p = 1$ and $p = \infty$ which are not spaces of type L_Φ^*.

[6] (Z), §§4.76, 4.77, 4.78.

[7] *A remark on conjugate series*, Proceedings of the London Mathematical Society, vol. 34 (1932), p. 396.

[8] (Z), §4.541.

[9] (Z), §7.31.

LEMMA B. *If $f(x)$ is absolutely continuous and has a derivative $f'(x) \in L^*_\Phi(0, 2\pi)$, then*

$$\|f(x+h) - f(x)\|_\Phi \leqq 2|h| \cdot \|f'\|_\Phi .$$

We have

$$\|f(t+h) - f(t)\| = |h| \cdot \|f'\| \left\{ \sup_g \left| \int_0^{2\pi} g(t) \left[\frac{f(t+h) - f(t)}{h\|f'\|} \right] dt \right| \right\}$$

$$\leqq |h| \cdot \|f'\| \left\{ \int_0^{2\pi} \Phi \left| \frac{f(t+h) - f(t)}{h\|f'\|} \right| dt + 1 \right\}.$$

We now need only show that

$$\int_0^{2\pi} \Phi \left| \frac{f(t+h) - f(t)}{h\|f'\|} \right| dt \leqq 1.$$

Let $e_h(t)$ be defined for each $h \neq 0$ such that $e_h(t) = 1$, $0 \leqq t \leqq |h|$, and zero elsewhere. Then

$$\Phi \left| \frac{f(x+h) - f(x)}{h\|f'\|} \right| = \Phi \left| \frac{1}{h} \int_0^h \frac{f'(t+x)}{\|f'\|} dt \right|$$

$$= \Phi \left| \frac{\displaystyle\int_0^{2\pi} e_h(t) \frac{f'(t+x)}{\|f'\|} dt}{\displaystyle\int_0^{2\pi} e_h(t)\, dt} \right|$$

$$\leqq \frac{1}{|h|} \int_0^{2\pi} e_h(t) \Phi \left| \frac{f'(t+x)}{\|f'\|} \right| dt$$

by Jensen's inequality.[10]

Consequently

$$\int_0^{2\pi} \Phi \left| \frac{f(x+h) - f(x)}{h\|f'\|} \right| dx \leqq \frac{1}{|h|} \int_0^{2\pi} e_h(t)\, dt \int_0^{2\pi} \Phi \left| \frac{f'(x+t)}{\|f'\|} \right| dx \leqq 1$$

since

$$\int_0^{2\pi} \Phi \left| \frac{f'(x+t)}{\|f'\|} \right| dx = \int_0^{2\pi} \Phi \left| \frac{f'(x)}{\|f'\|} \right| dx \leqq 1$$

because of the periodicity of $f'(x)$.

Let $\Omega(x)$ be a function not identically zero which satisfies the following conditions.

(i) $\Omega(x) \geqq 0$ and, at least for x greater than some x_0, decreases monotonically to zero as $x \to \infty$.

(ii) $\displaystyle\int_{x_0}^{\infty} \frac{\Omega(x)}{x} dx$ exists.

[10] (Z), §4.14.

We are now prepared to prove

THEOREM 1. *If the function $f(x)$ can be approximated, for each $n \geqq 1$, by a trigonometrical polynomial $T_n(x)$, of degree n at most, such that*

$$\| f - T_n \|_\Phi \leqq \frac{\Omega(n)}{n^r},$$

where r is a positive integer or zero, then $f(x)$ is equivalent to an absolutely continuous function having a derivative, $f^{(r)}(x)$, of order r, for which

$$\omega_\Phi(\delta; f^{(r)}) \leqq A\left[\delta \int_a^{a/\delta} \Omega(x)\,dx + \int_{1/\delta}^\infty \frac{\Omega(x)}{x}\,dx\right], \tag{1}$$

where a and A are constants which may depend on $f(x)$ but not on δ.

In the case $r = 0$, we understand $f^{(0)}(x) \equiv f(x)$. Two functions in this space are equivalent if they differ on at most a set of measure zero. The statement that $f(x)$ has a derivative $f^{(r)}(x)$ means that the derivatives $f^{(1)}(x)$, $f^{(2)}(x)$, $f^{(3)}(x), \cdots, f^{(r)}(x)$ exist on $(0, 2\pi)$, the last almost everywhere, and that $f(x)$, $f^{(1)}(x), f^{(2)}(x), \cdots, f^{(r-1)}(x)$ are absolutely continuous.

The method of proof is like that used by de la Vallée Poussin[11] for the case $p = \infty$.

Set

$$R_n(x) = f(x) - T_n(x),$$

$$\phi_k(x) = R_{a^k}(x) - R_{a^{k+1}}(x),$$

where a is an integer $\geqq 2$ such that $\Omega(x)$ is non-increasing for $x \geqq a$. Then $R_{a^2}(x) = f(x) - T_{a^2}(x)$. Since $T_{a^2}(x)$ is a trigonometrical polynomial, $T_{a^2}(x)$ and its derivatives of every order are in L_Φ^* and by Lemma B,

$$\| T_{a^2}^{(r)}(x + h) - T_{a^2}^{(r)}(x) \| \leqq 2|h| \cdot \| T_{a^2}^{(r+1)} \|,$$

so that $\omega_\Phi(\delta; T_{a^2}^{(r)}) \leqq M\delta$, where M is a constant independent of δ. The theorem will be proved if we can show that $R_{a^2}(x)$ satisfies the required conditions. For if $R(x)$ is absolutely continuous and equivalent to $R_{a^2}(x)$, then $R(x) + T_{a^2}(x)$ is absolutely continuous and equivalent to $f(x)$; moreover, if $R^{(r)}(x)$ exists and satisfies a relation of the type of (1), $f^{(r)} \equiv R^{(r)} + T_{a^2}^{(r)}$ will satisfy (1) since

$$\omega_\Phi(\delta; f^{(r)}) \leqq \omega_\Phi(\delta; R^{(r)}) + \omega_\Phi(\delta; T_{a^2}^{(r)})$$

and A in (1) may be chosen arbitrarily large.

Let $r \geqq 1$. Since $R_{a^2} = \sum_{k=2}^n \phi_k + R_{a^{n+1}}$,

$$\left\| R_{a^2} - \sum_{k=2}^n \phi_k \right\| \leqq \| R_{a^{n+1}} \| \leqq \frac{\Omega(a^{n+1})}{a^{(n+1)r}}.$$

[11] (P), §39.

This implies that $\sum_{k=2}^{\infty} \phi_k$ converges in L_Φ^* to R_{a^2}. By Lemma A,

$$(2) \qquad \|\phi_k'\| \leqq 2a^{k+1}\|\phi_k\| \leqq 2a^{k+1}(\|R_{a^k}\| + \|R_{a^{k+1}}\|) \leqq 4a\,\frac{\Omega(a^k)}{a^{k(r-1)}}$$

and consequently

$$(3) \qquad \sum_{k=2}^{\infty} \|\phi_k'\| \leqq \frac{4a^2}{a-1}\sum_{k=2}^{\infty}\frac{\Omega(a^k)(a^k - a^{k-1})}{a^{kr}} \leqq \frac{4a^2}{a-1}\int_a^\infty \frac{\Omega(x)}{x^r}\,dx < \infty.$$

This implies[12] that $\sum_{k=2}^{\infty} \phi_k'$ converges in L_Φ^* to a function in L_Φ^*, say $\rho'(x)$. Set

$$(4) \qquad R(x) = \int_{\bar{x}}^{x} \rho'(t) + R_{a^2}(\bar{x}),$$

where the point $\bar{x}$ is such that $\sum_{k=2}^{m_n} \phi_k(\bar{x}) \to R_{a^2}(\bar{x})$; this is possible since there must be a subsequence of $\left\{\sum_{k=2}^{m} \phi_k\right\}$ which converges to R_{a^2} almost everywhere. We have,[13] since

$$(5) \qquad \int_0^{2\pi}\left|\sum_{k=2}^{m_n}\phi_k' - \rho'\right|dt \leqq \left\|\sum_{k=2}^{m_n}\phi_k' - \rho'\right\| \to 0,$$

$$(6) \qquad \sum_{k=2}^{m_n}\int_{\bar{x}}^{x}\phi_k'(t)\,dt \to \int_{\bar{x}}^{x}\rho'(t)\,dt = R(x) - R_{a^2}(\bar{x}).$$

Thus from the equation

$$(7) \qquad \sum_{k=2}^{m_n}\phi_k(x) = \sum_{k=2}^{m_n}\int_{\bar{x}}^{x}\phi_k'(t)\,dt + \sum_{k=2}^{m_n}\phi_k(\bar{x}),$$

we have, by letting $n \to \infty$, $R_{a^2}(x) = R(x)$ almost everywhere.

Consider $r \geqq 2$. By Lemma A, corresponding to (3),

$$\sum_{k=2}^{\infty}\|\phi_k^{(2)}\| \leqq 2\sum_{k=2}^{\infty}a^{k+1}\|\phi_k'\| \leqq \frac{(2a)^3}{a-1}\int_a^\infty\frac{\Omega(x)}{x^{r-1}}\,dx < \infty;$$

this implies that $\sum_{k=2}^{\infty}\phi_k^{(2)}$ converges in L_Φ^* to a function in L_Φ^*, say $\rho^{(2)}(x)$. Proceeding as in the case of $\rho'(x)$, we set

$$(8) \qquad R'(x) = \int_{\bar{x}}^{x}\rho^{(2)}(t)\,dt + \rho'(\bar{x}),$$

where $\bar{x}$, $\{m_n\}$ are so chosen that $\sum_{k=2}^{m_n}\phi_k'(\bar{x}) \to \rho'(\bar{x})$. Then, by exactly the

[12] S. Banach, *Théorie des Opérations Linéaires*, Warsaw, 1932, p. 37.

[13] E. C. Titchmarsh, *The Theory of Functions*, Oxford, 1932, §12.53.

argument used for (4), we have $R'(x) = \rho'(x)$ almost everywhere. Thus $\sum_{k=2}^{\infty} \phi_k'$ converges in L_Φ^* to R' and we can denote the derivative of R by R' instead of ρ'.

For $r \geqq 1$ we have shown the existence of an absolutely continuous function $R(x)$ equivalent to $R_{a^2}(x)$ and having a derivative $R'(x)$ which, for $r \geqq 2$, is in turn absolutely continuous.

The existence of $R^{(r)}$ now follows by induction. Let $s < r$, and suppose $R^{(s-1)}$ exists, is absolutely continuous, and has as its derivative $\rho^{(s)}(x)$ defined as the limit in L_Φ^* of $\sum_{k=2}^{\infty} \phi_k^{(s)}$. This hypothesis implies the existence of an absolutely continuous function, $R^{(s)}(x)$, equivalent to $\rho^{(s)}(x)$ which has as its derivative the function $\rho^{(s+1)}(x)$ defined as the limit in L_Φ^* of $\sum_{k=2}^{\infty} \phi_k^{(s+1)}(x)$. The proof is as follows. By Lemma A

$$(9) \qquad \| \phi_k^{(s+1)} \| \leqq 2a^{k+1} \| \phi_k^{(s)} \| \leqq \cdots \leqq (2a^{k+1})^{s+1} \| \phi_k \|,$$

and, corresponding to (3), we have

$$(10) \qquad \sum_{k=2}^{\infty} \| \phi_k^{(s+1)} \| \leqq 2^{s+2} a^{s+1} \sum_{k=2}^{\infty} \frac{\Omega(a^k)}{a^{k(r-s-1)}} \leqq \frac{(2a)^{s+2}}{a-1} \int_a^{\infty} \frac{\Omega(x)\,dx}{x^{r-s}} < \infty .$$

This implies that $\sum_{k=2}^{\infty} \phi_k^{(s+1)}(x)$ converges in L_Φ^* to a function in L_Φ^*, say $\rho^{(s+1)}(x)$. Corresponding to (4) and (8) we set

$$(11) \qquad R^{(s)}(x) = \int_{\bar{x}}^{x} \rho^{(s+1)}(t)\,dt + \rho^{(s)}(\bar{x}),$$

where $\bar{x}$, $\{m_n\}$ are so chosen that $\sum_{k=2}^{m_n} \phi_k^{(s)}(\bar{x}) \to \rho^{(s)}(\bar{x})$. By the argument used in cases $s = 0$ and $s = 1$ with equations corresponding to (5), (6), (7) we have $R^{(s)}(x) = \rho^{(s)}(x)$ almost everywhere.

When $s = r - 1$, we denote $\rho^{(s+1)}(x)$ by $R^{(r)}(x)$. For $r = 0$, $R^{(0)}(x) = R_{a^2}(x)$, $\phi_k^{(0)} = \phi_k$. We must show that $R^{(r)}(x)$ satisfies (1). Since, for $r \geqq 0$, $\sum_{k=2}^{m} \phi_k^{(r)} \to R^{(r)}$ in L_Φ^*, we have from Lemma B and (9),

$$\begin{aligned} \| R^{(r)}(x+h) - R^{(r)}(x) \| &\leqq \sum_{k=2}^{m} \| \phi_k^{(r)}(x+h) - \phi^{(r)}(x) \| \\ &\qquad + \sum_{m+1}^{\infty} \| \phi_k^{(r)}(x+h) - \phi_k^{(r)}(x) \| \\ &\leqq 2|h| \sum_{k=2}^{m} \| \phi_k^{(r+1)} \| + 2 \sum_{m+1}^{\infty} \| \phi_k^{(r)} \| \\ &\leqq 2^{r+2} |h| \sum_{k=2}^{m} a^{(k+1)(r+1)} \| \phi_k \| + 2^{r+1} \sum_{m+1}^{\infty} a^{(k+1)r} \| \phi_k \| \end{aligned}$$

$$\leqq 2^{r+3}|h|a^{r+2}\sum_{k=2}^{m}a^{k-1}\Omega(a^k)+2^{r+2}a^r\sum_{m+1}^{\infty}\Omega(a^k)$$

$$\leqq \frac{2(2a)^{r+2}}{a-1}|h|\int_a^{a^m}\Omega(x)\,dx+\frac{2(2a)^{r+1}}{a-1}\int_{a^m}^{\infty}\frac{\Omega(x)}{x}\,dx.$$

If m is now chosen subject to the conditions $a^{m-1}\leqq\delta^{-1}<a^m$ and $A\geqq 2(2a)^{r+2}(a-1)^{-1}$, we have

$$\omega_\Phi(\delta;R^{(r)})\leqq A\left[|\delta|\int_a^{a/\delta}\Omega(x)\,dx+\int_{1/\delta}^{\infty}\frac{\Omega(x)}{x}\,dx\right].$$

If the $\Omega(x)$ considered in Theorem 1 satisfies only the hypothesis (i), namely, that $\Omega(x)\downarrow 0$, and not (ii) also, then the condition $\|f-T_n\|\leqq\Omega(n)n^{-r}$ does not necessarily imply that (1) holds. In fact $f^{(r)}(x)$ may not even ϵL_Φ^*. For consider the space L_2 and let

$$f(x)=\sum_{k=1}^{\infty}\Omega(k)k^{-\frac{3}{2}}\sin kx.$$

Then, if $s_n=s_n(x;f)$ is the n-th partial sum of the Fourier series of $f(x)$,

$$\|f-s_n\|_2=\left[\sum_{k=n+1}^{\infty}\left(\frac{\Omega(k)}{k^{\frac{3}{2}}}\right)^2\right]^{\frac{1}{2}}\leqq\Omega(n)\left(\sum_{k=n+1}^{\infty}\frac{1}{k^3}\right)^{\frac{1}{2}}\leqq\frac{\Omega(n)}{n}.$$

But $f'(x)=\sum_{k=1}^{\infty}\Omega(k)k^{-\frac{1}{2}}\cos kx$ may not even be in L_2 since

$$\|f'\|_2=\left[\sum_{k=1}^{\infty}\left(\frac{\Omega(k)}{k^{\frac{1}{2}}}\right)^2\right]^{\frac{1}{2}}$$

does not necessarily converge; for example, if $\Omega(x)=[\log(x+1)]^{-\frac{1}{2}}$.

We remark that the theorem just proved holds for L_1 as well as for L_Φ^* (and thus for L_p, $p>1$). The proof for L_1 is the above proof with the L_Φ^* norm replaced by the L_1 norm since the inequalities of Lemmas A and B hold also for L_1.

As an immediate corollary of Theorem 1 we have the positive assertion of the following theorem.

THEOREM 2. *If the function $f(x)$ can be approximated for each $n\geqq 1$, by a trigonometrical polynomial, $t_n(x)$, of degree n at most, such that $\|f-t_n\|_p=O(n^{-\alpha})$, $p\geqq 1$, then*

(i) *if* $0<\alpha<1$, $f(x)\,\epsilon\,\mathrm{Lip}(\alpha,p)$;

(ii) *if* $\alpha=1$, $\omega_p(\delta;f)=O\left(\delta\log\frac{1}{\delta}\right)$.

Moreover there exist functions for which $\|f-t_n\|_p=O(n^{-1})$ which do not belong to Lip$(1,p)$.

In the previous theorem we choose $r = 0$, $L^*_\Phi \equiv L_p$, $\Omega(x) = Mx^{-\alpha}$, where M is a positive constant. Then

$$\omega_p(\delta; f) = O\left\{\delta \int_a^{a/\delta} x^{-\alpha}\,dx + \int_{1/\delta}^{\infty} x^{-1-\alpha}\,dx\right\}.$$

This gives

(i) $\omega_p(\delta;f) = O(\delta^\alpha), 0 < \alpha < 1$;

(ii) $\omega_p(\delta;f) = O\left(\delta \log \dfrac{1}{\delta}\right), \quad \alpha = 1.$

If $f(x) \in \mathrm{Lip}(\alpha, p)$, $\alpha = 1$, $p > 1$, then $f(x)$ is equivalent[14] to the indefinite integral of a function in L_p. Consider the function $f(x) = \sum_{n=1}^{\infty} n^{-\frac{3}{2}} e^{inx}$. We have

$$\|f - s_n\|_2 = \left(\sum_{\nu=n+1}^{\infty} \nu^{-3}\right)^{\frac{1}{2}} \backsimeq 2^{-\frac{1}{2}} n^{-1}.$$

But $f(x)$ is not in Lip(1, 2) since

$$f'(x) \backsim i \sum_{\nu=1}^{\infty} \nu^{-\frac{1}{2}} e^{i\nu x}$$

and $f'(x)$ is not in L_2. Indeed, for $h > 0$,

$$\|f(x+h) - f(x-h)\|_2 = \left(\sum_{n=1}^{\infty} n^{-3} \sin^2 nh\right)^{\frac{1}{2}} = h\left(\sum_{n=1}^{\infty} \frac{1}{n}\left[\frac{\sin nh}{nh}\right]^2\right)^{\frac{1}{2}}$$
$$\geqq \left(\frac{2}{\pi}\right)^2 h\left(\sum_{n=1}^{[\pi/2h]} n^{-1}\right)^{\frac{1}{2}} \geqq \left(\frac{2}{\pi}\right)^2 h\left(\log\left[\frac{\pi}{2h}\right]\right)^{\frac{1}{2}},$$

so that

$$\|f(x+h) - f(x-h)\|_2 \neq o\left(h\left[\log \frac{1}{h}\right]^{\frac{1}{2}}\right).$$

If $\alpha > 1$, we may take $\Omega(x) = x^{r-\alpha}$, where r is an integer such that $0 < \alpha - r \leqq 1$. Then $f^{(r)}(x)$ exists and the conclusion of the theorem holds with $f^{(r)}(x)$ in place of $f(x)$.

We now turn to the consideration of theorems of the converse type.

Theorem 3. *If $f(x) \in L^*_\Phi$ possesses a derivative of order r, say $f^{(r)}(x)$, in L^*_Φ, where r is a positive integer or zero, then, for any positive integer n, $f(x)$ may be approximated in L^*_Φ by a trigonometrical polynomial $t_n(x)$, of order n at most, such that*

$$\|f - t_n\|_\Phi = O\left(n^{-r}\omega_\Phi\left(\frac{1}{n}; f^{(r)}\right)\right).$$

[14] (Z), §4.7, (8). By the use of lacunary series and the inequalities of (Z), §§9.601 and 9.602, we need not restrict the example to the space L_2.

To prove this theorem we use the method and notation of de la Vallée Poussin.[15] We set

$$\lambda = \left[\frac{r}{2}\right] \quad \text{and} \quad \phi(t) = \sum_{k=0}^{\lambda} a_k f\left(x + \frac{2t}{2^k}\right);$$

the $\lambda + 1$ constants a_k, $k = 0, 1, 2, \cdots, \lambda$ being so determined that $\phi(0) = f(x)$, $\phi^{(2s)}(0) = 0$, $s = 1, 2, \cdots, \lambda$. The trigonometrical polynomial $t_n(x)$ is defined by the equation

$$t_n(x) = \frac{1}{\tau(\lambda + 2)} \int_{-\infty}^{+\infty} \phi\left(\frac{t}{m}\right)\left(\frac{\sin t}{t}\right)^{2\lambda+4} dt,$$

where

$$\tau(\lambda + 2) = \int_{-\infty}^{+\infty} \left(\frac{\sin t}{t}\right)^{2\lambda+4} dt.$$

The order n of $t_n(x)$ is $(\lambda + 2)2^{\lambda} m - 1$. Since $\phi(0) = f(x)$, we write

$$t_n(x) - f(x) = \frac{1}{\tau(\lambda + 2)} \int_0^{\infty} F\left(\frac{t}{m}\right)\left(\frac{\sin t}{t}\right)^{2\lambda+4} dt,$$

where $F(t) = \phi(t) + \phi(-t) - 2\phi(0)$.

With these definitions

$$\phi^{(r)}(t) = \sum_{k=0}^{\lambda} a_k \left(\frac{1}{2^{(k-1)r}}\right) f^{(r)}\left(x + \frac{2t}{2^k}\right)$$

and

$$\begin{aligned} F^{(r)}(t) &= \phi^{(r)}(t) + \phi^{(r)}(-t) - 2\phi^{(r)}(0), \quad r \text{ even}; \\ &= \phi^{(r)}(t) - \phi^{(r)}(-t), \quad r \text{ odd}. \end{aligned}$$

Consequently,

$$\begin{aligned} \| F^{(r)}(t) \| &\leqq \sum_{k=0}^{\lambda} \frac{|a_k|}{2^{(k-1)r}} \left\| f^{(r)}\left(x + \frac{2t}{2^k}\right) + f^{(r)}\left(x - \frac{2t}{2^k}\right) - 2f^{(r)}(x) \right\|, \quad r \text{ even}, \\ &\leqq \sum_{k=0}^{\lambda} \frac{|a_k|}{2^{(k-1)r}} \left\| f^{(r)}\left(x + \frac{2t}{2^k}\right) - f^{(r)}(x) + f^{(r)}(x) - f^{(r)}\left(x - \frac{2t}{2^k}\right) \right\|, \quad r \text{ odd}, \end{aligned}$$

or, finally,

$$\| F^{(r)}(t) \| = O(\omega_\Phi(2t; f^{(r)})).$$

We also have, for $r \geqq 1$,

$$F(t) = \int_0^t \int_0^{t_1} \cdots \int_0^{t_{r-1}} F^{(r)}(u)\, du\, dt_{r-1} \cdots dt_2\, dt_1$$

[15] (P), pp. 47–50.

so that

$$\| F(t) \| \leqq \int_0^t \int_0^{t_1} \cdots \int_0^{t_{r-1}} \| F^{(r)}(u) \| \, du dt_{r-1} \cdots dt_2 dt_1$$

$$= O\left[\int_0^t \int_0^{t_1} \cdots \int_0^{t_{r-1}} \omega_\Phi(2u; f^{(r)}) \, du dt_{r-1} \cdots dt_2 dt_1\right].$$

Thus

$$\left\| F\left(\frac{t}{m}\right) \right\| = O\left[\frac{1}{m^r} \int_0^t \int_0^{t_1} \cdots \int_0^{t_{r-1}} \omega_\Phi\left(\frac{2u}{m}; f^{(r)}\right) du dt_{r-1} \cdots dt_2 dt_1\right].$$

Since $f^{(r)}(x)$ is periodic,

$$\omega_\Phi\left(\frac{2u}{m}; f^{(r)}\right) \leqq (2u + 1)\, \omega_\Phi\left(\frac{1}{m}; f^{(r)}\right).$$

Thus we have

$$\left\| F\left(\frac{t}{m}\right) \right\| = O\left\{\frac{1}{m^r} \omega_\Phi\left(\frac{1}{m}; f^{(r)}\right) \int_0^t \int_0^{t_1} \cdots \int_0^{t_{r-1}} (2u + 1) \, du dt_{r-1} \cdots dt_2 dt_1\right\}$$

$$= O\left\{m^{-r} t^r \omega_\Phi\left(\frac{1}{m}; f^{(r)}\right)\right\}.$$

Now

$$\| t_n - f \| \leqq \frac{1}{\tau(\lambda + 2)} \int_0^\infty \left\| F\left(\frac{t}{m}\right) \right\| \left(\frac{\sin t}{t}\right)^{2\lambda+4} dt$$

$$= O\left\{m^{-r} \omega_\Phi\left(\frac{1}{m}; f^{(r)}\right) \int_0^\infty t^r \left(\frac{\sin t}{t}\right)^{2\lambda+4} dt\right\}$$

$$= O\left\{m^{-r} \omega_\Phi\left(\frac{1}{m}; f^{(r)}\right)\right\}.$$

Since $m = (n + 1)(\lambda + 2)^{-1} 2^{-\lambda}$, we have

$$\| t_n - f \| = O\left\{n^{-r} \omega_\Phi\left(\frac{1}{n}; f^{(r)}\right)\right\}.$$

THEOREM 4. *If $f(x) \in \text{Lip}(\alpha, p)$, $p \geqq 1$, $0 < \alpha \leqq 1$, then, for any positive integer n, $f(x)$ may be approximated in L_p by a trigonometrical polynomial, $t_n(x)$, of order n such that*

$$\| f - t_n \|_p = O(n^{-\alpha}).$$

We put $r = 0$, $\lambda = 0$, and $\omega_p(n^{-1}; f^{(r)}) = Mn^{-\alpha}$ in Theorem 3. For $p > 1$, the result is obtained by taking L_Φ^* to be L_p. For $p = 1$, the result is obtained by carrying out the proof with L_1 in place of L_Φ^*.

Theorems 2 and 4 give the proof and range of Hardy and Littlewood's theorem. The question concerning the remark following the theorem is answered by Theorem 5.

In the following $s_n = s_n(f) = s_n(x; f)$ denotes the n-th partial sum of the Fourier series of $f(x)$ and $\sigma_n = \sigma_n(x; f)$ denotes the n-th $(C, 1)$ mean of s_n.

LEMMA C. *If $f(x) \in L_p$ and $t_n(x)$ is an arbitrary trigonometric polynomial of degree $n \geqq 1$ at most, then*

(i) *if* $p > 1$, $\| f - s_n \|_p \leqq A \| f - t_n \|_p$,

(ii) *if* $p = 1$, $\| f - s_n \|_1 \leqq A(1 + \log n) \| f - t_n \|_1$,

where A is independent of $f(x)$ and n.

We may write

$$\| f - s_n \| = \| f - t_n + t_n - s_n \| \leqq \| f - t_n \| + \| s_n - t_n \|.$$

The trigonometric polynomial $s_n(x; f) - t_n(x) = s_n(x; f - t_n)$. Hence for $p > 1$, by an inequality of M. Riesz,[16] we have

$$\| s_n - t_n \| \leqq A \| f - t_n \|, \qquad p > 1.$$

When $p = 1$, we have

$$\| s_n - t_n \| = \frac{1}{\pi} \int_0^{2\pi} \left| \int_0^{2\pi} [f(x + u) - t_n(x + u)] D_n(u) \, du \right| dx,$$

where $D_n(u)$ is the Dirichlet kernel. Interchanging the order of integration, we have

$$\| s_n - t_n \| \leqq \frac{1}{\pi} \| f - t_n \| \int_0^{2\pi} | D_n(u) | \, du \leqq A(1 + \log n) \| f - t_n \|$$

since[17]

$$\frac{1}{\pi} \int_0^{2\pi} | D_n(u) | \, du \backsimeq \frac{4}{\pi^2} \log n.$$

THEOREM 5. *If $f(x) \in \mathrm{Lip}(\alpha, p)$, $0 < \alpha \leqq 1$, then*

(i) *when* $p > 1$, $\| f - s_n \|_p = O(n^{-\alpha})$;

(ii) *when* $p = 1$, $\| f - s_n \|_1 = O(n^{-\alpha} \log n)$.

For $p > 1$, $\alpha < 1$, O-large cannot be replaced by o-small and there exist functions in $\mathrm{Lip}(1, 1)$ *for which* $\| f - s_n \| \neq o(n^{-1} \log n)$.

The positive assertion of this theorem is a corollary of Theorem 4 by application of Lemma C.

To show that O-large cannot be replaced by o-small for $p > 1$, $\alpha < 1$ consider the function

$$f(x) = \sum_{m=1}^{\infty} \frac{\cos 2^m x}{2^{m\alpha}}, \quad 0 < \alpha < 1.$$

[16] (Z), §7.3, (1). Since $\| s_n \| \leq \| s_n^* \| + \| f \|$, $p > 1$, we have $\| s_n \| \leq (2A_p + 1) \| f \|$.

[17] (Z), §8.3.

This function[18] is in Lip(α) for each α and, a fortiori, in Lip(α, p) for each α and all p. We have,[19] however,

$$\|f - s_{2^n}\| \geqq B\left(\sum_{\nu=n+1}^{\infty} \frac{1}{2^{2\nu\alpha}}\right)^{\frac{1}{2}} \neq o(2^{-n\alpha}).$$

The function

$$f(x) = \sum_{\nu=1}^{\infty} \frac{\sin \nu x}{\nu}$$

is in the class Lip(1, 1). Making a slight change in notation, we set

$$t_n(x) = t_n(x;f) = h_m \int_{-\pi}^{\pi} f(x+t) \left(\frac{\sin \frac{1}{2}mt}{m \sin \frac{1}{2}t}\right)^4 dt,$$

where $n = 2m - 2$ and

$$\frac{1}{h_m} = \int_{-\pi}^{\pi} \left(\frac{\sin \frac{1}{2}mt}{m \sin \frac{1}{2}t}\right)^4 dt.$$

Our definition[20] of $t_n(x)$ differs slightly from that of Theorem 3.

For

$$f(x) = \sum_{\nu=1}^{\infty} \frac{\sin \nu x}{\nu},$$

we have

$$t_n'(x;f) = \pi h_m \left(\frac{\sin \frac{1}{2}mx}{m \sin \frac{1}{2}x}\right)^4 - \tfrac{1}{2}.$$

To show this we write

$$\begin{aligned} t_n(x;f) &= h_m \int_{-\pi}^{\pi} \sum_{\nu=1}^{\infty} \frac{\sin \nu(x+t)}{\nu} \left(\frac{\sin \frac{1}{2}mt}{m \sin \frac{1}{2}t}\right)^4 dt \\ &= h_m \sum_{\nu=1}^{n} \frac{\sin \nu x}{\nu} \int_{-\pi}^{\pi} \cos \nu t \left(\frac{\sin \frac{1}{2}mt}{m \sin \frac{1}{2}t}\right)^4 dt. \end{aligned}$$

Then

$$t_n'(x;f) = h_m \sum_{\nu=1}^{n} \cos \nu x \int_{-\pi}^{\pi} \cos \nu t \left(\frac{\sin \frac{1}{2}mt}{m \sin \frac{1}{2}t}\right)^4 dt.$$

[18] (Z), §2.9, (3).

[19] (Z), §9.602, (1).

[20] We are here using the notation of Dunham Jackson, (D), page 3, rather than that of de la Vallée Poussin. However

$$\begin{aligned} \|t_n - f\|_1 &= \left\| h_m \int_{-\pi}^{\pi} [f(x+t) - f(x)] \left(\frac{\sin \frac{1}{2}mt}{m \sin \frac{1}{2}t}\right)^4 dt \right\|_1 \\ &= O\left(m \int_{-\pi}^{\pi} |t| \left(\frac{\sin \frac{1}{2}mt}{m \sin \frac{1}{2}t}\right)^4 dt\right) = O\left(\frac{1}{n}\right). \end{aligned}$$

This equation implies

$$t_n'(x;f) = \pi h_m \left(\frac{\sin \frac{1}{2}mx}{m \sin \frac{1}{2}x}\right)^4 - \tfrac{1}{2}.$$

Now

$$s_n'(x;f) = \sum_{\nu=1}^{n} \cos \nu x = D_n(x) - \tfrac{1}{2},$$

where $D_n(x)$ is the Dirichlet kernel.

By an inequality of F. Riesz,[21]

$$\| t_n - s_n \| \geqq \frac{1}{n} \| t_n' - s_n' \| = \frac{1}{n} \int_{-\pi}^{\pi} \left| D_n(u) - \pi h_m \left(\frac{\sin \frac{1}{2}mu}{m \sin \frac{1}{2}u}\right)^4 \right| du$$

$$\geqq \frac{1}{n} \int_{-\pi}^{\pi} | D_n(u) | \, du - \frac{\pi h_m}{n} \int_{-\pi}^{\pi} \left(\frac{\sin \frac{1}{2}mu}{m \sin \frac{1}{2}u}\right)^4 du \geqq A \frac{\log n}{n} - \frac{\pi}{n}.$$

This gives $\| t_n - s_n \| \neq o(n^{-1} \log n)$. Since

$$\| f - t_n \| + \| f - s_n \| \geqq \| t_n - s_n \| \qquad \text{and} \qquad \| f - t_n \| = O(n^{-1}),$$

we have $\| f - s_n \| \neq o(n^{-1} \log n)$.

THEOREM 6. *If* $f(x) \in \mathrm{Lip}(\alpha, p)$, $0 < \alpha \leqq 1$, *then*

(i) *if* $p > 1$ *or if* $p = 1$, $\alpha < 1$, $\| f - \sigma_n \| = O(n^{-\alpha})$;

(ii) *if* $p = \alpha = 1$, $\| f - \sigma_n \| = O\left(\frac{\log n}{n}\right)$.

For the case (i), *O-large cannot be replaced by o-small.*

For the cases[22] $p \geqq 1$, $\alpha < 1$, and $p = \alpha = 1$ we write

$$\| \sigma_n - f \| = \frac{1}{\pi} \int_0^{2\pi} \left| \int_0^{2\pi} [f(x + t) - f(x)] K_n(t) \, dt \right| dx,$$

where $K_n(t)$ is the Fejér kernel. The result follows by Minkowski's inequality[23] since

$$\int_0^{2\pi} t^\alpha K_n(t) \, dt = O(n^{-\alpha}), \quad \alpha < 1;$$

$$= O(n^{-1} \log n), \quad \alpha = 1.$$

When $p > 1$, $f(x) \in \mathrm{Lip}(1, p)$ is equivalent to the indefinite integral of a function in L_p. Since

$$\| \sigma_n - s_n \| = \frac{1}{n+1} \| \tilde{s}_n' \| \leqq \frac{2A_p + 1}{n + 1} \| f' \| = O(n^{-1}),$$

[21] (Z), §7.31, (b).

[22] The case $p > 1$ and $\alpha < 1$ was obtained by O. Szász, *Über die Fourierschen Reihen gewisser Funktionenklassen*, Mathematische Annalen, vol. 100 (1928), pp. 530–536.

[23] (Z), §4.13, (4).

where $\bar{s}'_n(x)$ is the n-th partial sum of the conjugate derived series of $f(x)$, the result for $p > 1$, $\alpha = 1$ follows from Theorem 5 and the inequality

$$\| f - \sigma_n \| \leqq \| f - s_n \| + \| s_n - \sigma_n \|.$$

To show that O-large cannot be replaced by o-small in (i) we apply Lemma C to Theorem 5 for the case $p > 1$, $\alpha < 1$. For $p > 1$, $\alpha = 1$ consider $f(x) = \cos x$. This function belongs to Lip(1, p) for every p. But

$$\| f - \sigma_n \| = \left\| f - \frac{n-1}{n} f \right\| = \frac{1}{n} \| f \| \neq o\left(\frac{1}{n}\right).$$

Corresponding to the case[24] $p = \infty$, if $f(x)$ can be approximated in L_p, $p \geqq 1$ by $s_n(x; f)$ such that the order of the approximation is $\omega(n)$, it follows from Lemma C that no approximation of $f(x)$ by a trigonometrical polynomial of order n can, for $p > 1$, be $o[\omega(n)]$, and, for $p = 1$, $o[\omega(n)(1 + \log n)^{-1}]$.

Also corresponding to the case[25] $p = \infty$ we have

THEOREM 7. *A necessary and sufficient condition that a function $f(x)$ periodic in 2π belong to the class* Lip(α, p), $0 < \alpha \leqq 1$, $p \geqq 1$, *is that $\sigma_n(x; f)$ belong to* Lip(α, p) *uniformly in n.*

The necessity follows immediately from the inequality $\| \sigma_n \| \leqq \| f \|$ and the sufficiency from the same inequality on the application of the Fatou lemma.

THEOREM 8. *Let $F(x)$ be periodic in 2π and the indefinite integral of a function $f(x) \in L_p$, $p \geqq 1$. Then*

(i) $\| F - s_n(F) \| \leqq \dfrac{A}{n+1} \| f - s_n(f) \|, \quad p > 1;$

(ii) $\| F - s_n(F) \| \leqq A\left(\dfrac{1 + \log n}{n}\right) \| f - s_n(f) \|, \quad p = 1.$

Let

$$f(x) \sim \sum_{\nu=1}^{\infty} a_\nu \cos \nu x + b_\nu \sin \nu x$$

(in order that $F(x)$ be periodic $a_0 = 0$) and let $\bar{f}(x)$ denote the conjugate function. Choose $m > k > n$. We may write

$$\begin{aligned} s_k(x; F) - s_n(x; F) &= \sum_{n+1}^{k} \frac{1}{\nu} (a_\nu \sin \nu x - b_\nu \cos \nu x) \\ &= \frac{1}{n+1} [s_m(x; \bar{f}) - s_n(x; \bar{f})] \\ &\quad - \sum_{n+1}^{k-1} \frac{1}{\nu(\nu+1)} [s_m(x; \bar{f}) - s_\nu(x; \bar{f})] - \frac{1}{k} [s_m(x; \bar{f}) - s_k(x; \bar{f})]. \end{aligned}$$

[24] (P), p. 22.
[25] (Z), §4.719.

Then

$$\| s_k(F) - s_n(F) \| \leqq \frac{1}{n+1} \| s_m(\tilde{f}) - s_n(\tilde{f}) \|$$

$$+ \sum_{n+1}^{k-1} \frac{1}{\nu(\nu+1)} \| s_m(\tilde{f}) - s_\nu(\tilde{f}) \| + \frac{1}{k} \| s_m(\tilde{f}) - s_k(\tilde{f}) \|.$$

First let $m \to \infty$ and then $k \to \infty$. We obtain,[26] for $p > 1$,

$$\| F - s_n(F) \| \leqq \frac{1}{n+1} \| \tilde{f} - s_n(\tilde{f}) \| + \sum_{n+1}^{\infty} \frac{1}{\nu(\nu+1)} \| \tilde{f} - s_\nu(\tilde{f}) \|$$

$$\leqq \frac{A_p}{n+1} \| f - s_n(f) \| + A_p \sum_{n+1}^{\infty} \frac{1}{\nu(\nu+1)} \| f - s_\nu(f) \|,$$

and (i) follows since, for $\nu > n$, $\| f - s_\nu(f) \| \leqq M \| f - s_n(f) \|$.

The above proof is due to the referee; it eliminates the $\log n$ from the first part of the theorem. For $p = 1$

$$F(x) - s_n(x; F) = \sum_{\nu=1}^{\infty} \frac{1}{\nu} (a_\nu \sin \nu x - b_\nu \cos \nu x)$$

$$= \frac{-1}{\pi} \int_{-\pi}^{\pi} \{ f(x+t) - s_n(x+t; f) \} \left(\sum_{n+1}^{\infty} \frac{\sin \nu t}{\nu} \right) dt.$$

The interchange of summation and integration is possible since $f(x) \in L$ and $\sum_{n+1}^{\infty} \nu^{-1} \sin \nu x$ is boundedly convergent. We have

$$\| F - s_n(F) \| \leqq \frac{1}{\pi} \int_{-\pi}^{\pi} \left(\int_{-\pi}^{\pi} | f(x+t) - s_n(x+t; f) | \, dx \right) \left| \sum_{n+1}^{\infty} \frac{\sin \nu t}{\nu} \right| dt$$

$$\leqq \frac{1}{\pi} \| f - s_n(f) \| \int_{-\pi}^{\pi} \left| \sum_{n+1}^{\infty} \frac{\sin \nu t}{\nu} \right| dt.$$

By Theorem 5

$$\int_{-\pi}^{\pi} \left| \sum_{n+1}^{\infty} \frac{\sin \nu t}{\nu} \right| dt \leqq A \left(\frac{1 + \log n}{n} \right)$$

and the theorem follows.

UNIVERSITY OF FLORIDA.

[26] (Z), §§7.21, 7.3.

A NOTE ON NON-ASSOCIATIVE ALGEBRAS

By N. Jacobson

It is the purpose of this note to obtain relations between an arbitrary algebra $\mathfrak{R}$ (not necessarily associative) and an algebra $\mathfrak{A}$ (necessarily associative) of linear transformations determined by $\mathfrak{R}$. If $\mathfrak{A}$ is simple, the centrum $\mathfrak{C}$ of $\mathfrak{A}$ is an algebraic field and $\mathfrak{R}$ may be regarded as an algebra over $\mathfrak{C}$. When this is done $\mathfrak{R}$ becomes a *normal simple algebra*, i.e., remains simple when this field is extended to its algebraic closure. A field having this property for algebras of characteristic 0 has been defined previously but less directly by Landherr.[1] Some of our results have been announced for Lie algebras of characteristic 0 by Albert.[2]

1. Let $\mathfrak{R}$ be an arbitrary algebra (not necessarily associative or commutative) with a finite basis over a commutative field Φ; $\mathfrak{R}$ is a finite dimensional vector space over Φ in which there is defined a composition xy of pairs of elements x, y such that

$$(1) \qquad (x + y)z = xz + yz, \qquad z(x + y) = zx + zy,$$

$$(2) \qquad (xy)\alpha = x(y\alpha) = (x\alpha)y, \qquad \alpha \in \Phi.$$

The mapping $x \rightarrow xa \equiv xA_r$ of $\mathfrak{R}$ on itself will be called the *right multiplication* determined by a. Equations (1) and (2) show that A_r is a linear transformation in the vector space $\mathfrak{R}$. Similarly we define the *left multiplication* determined by a as $x \rightarrow ax \equiv xA_l$. Let $\mathfrak{A}$ be the enveloping algebra of the left and right multiplications of $\mathfrak{R}$, i.e., the smallest algebra of linear transformations in $\mathfrak{R}$ containing all the left and right multiplications. The elements of $\mathfrak{A}$ are sums of terms of the type $A_{1i_1} \cdots A_{si_s}$ ($i_\alpha = r$ or l) where A_{ji_j} is a multiplication determined by a_j. We shall therefore denote an arbitrary element of $\mathfrak{A}$ by $\Sigma A_{1i_1} \cdots A_{si_s}$ (not summed on i_α!). Thus $\mathfrak{A}$ may also be defined as the smallest ring of linear transformations containing all the multiplications.

If $a_1, \cdots, a_n$ is a basis for $\mathfrak{R}$ over Φ and A is a linear transformation in this vector space, then A is completely determined by the matrix (α_{ij}) such that $a_j A = \Sigma a_i \alpha_{ij}$. The correspondence between A and the matrix (α_{ij}) determines, as is well known, a reciprocal isomorphism between the ring of all linear transformations in $\mathfrak{R}$ over Φ and the matrix ring Φ_n of all n-rowed square matrices

Received February 11, 1937; presented to the American Mathematical Society, April 9, 1937. The author is a National Research Fellow.

[1] W. Landherr, *Über einfache Liesche Ringe*, Hamb. Abhandlungen, vol. 11 (1935), pp. 41–64.

[2] Bull. Am. Math. Soc., vol. 41 (1935), p. 344.

with coördinates in Φ. In particular $\mathfrak{A}$ may be represented (reciprocally) as a subring of Φ_n.

If Σ is an extension of the field Φ, we define $\mathfrak{R}_\Sigma$ to be the algebra over Σ having the same basis as $\mathfrak{R}$ has over Φ, i.e., if $\mathfrak{R} = a_1\Phi + \cdots + a_n\Phi$, then $\mathfrak{R}_\Sigma = a_1\Sigma + \cdots + a_n\Sigma$. Any linear transformation A in $\mathfrak{R}$ over Φ has a unique extension to a linear transformation in $\mathfrak{R}_\Sigma$ over Σ. The matrix of A (relative to the basis $a_1, \cdots, a_n$) and of its extension are, of course, identical. It follows readily from the matrix representation that the enveloping algebra of the multiplications of $\mathfrak{R}_\Sigma$ is the extension algebra $\mathfrak{A}_\Sigma$.

As usual we define a (two-sided) *ideal* $\mathfrak{S}$ of $\mathfrak{R}$ as a subspace of $\mathfrak{S}$ such that $\mathfrak{S} \supset zx, xz$ for all $z \in \mathfrak{S}$ and $x \in \mathfrak{R}$. Thus $\mathfrak{S}$ is a subspace invariant under all the left and right multiplications and hence under all the transformations of $\mathfrak{A}$. $\mathfrak{R}$ is a *direct sum* of the ideals $\mathfrak{R}_1, \cdots, \mathfrak{R}_k$ ($\mathfrak{R} = \mathfrak{R}_1 \oplus \cdots \oplus \mathfrak{R}_k$) if every x in $\mathfrak{R}$ is expressible uniquely as $x_1 + \cdots + x_k$, x_i in $\mathfrak{R}_i$. This notion coincides with that of decomposability of $\mathfrak{R}$ relative to the system $\mathfrak{A}$. Since $x_ix_j \in$ the intersection $\mathfrak{R}_i \cap \mathfrak{R}_j = 0$, $x_ix_j = 0$ for any $x_i \in \mathfrak{R}_i$, $x_j \in \mathfrak{R}_j$. $\mathfrak{R}$ is *simple* if it has no proper ideal, or in other words, if $\mathfrak{A}$ is an irreducible system of linear transformations.[3]

THEOREM 1. *A necessary and sufficient condition that $\mathfrak{R}$ be a direct sum of simple algebras is that $\mathfrak{A}$ be a completely reducible system.*

If $\mathfrak{R} = \mathfrak{R}_1 \oplus \cdots \oplus \mathfrak{R}_k$, where the $\mathfrak{R}_i$ are simple, then the $\mathfrak{R}_i$ are irreducible subspaces and $\mathfrak{A}$ is completely reducible. Conversely if the $\mathfrak{R}_i$ are irreducible, they are simple. For let $\mathfrak{S}_i$ be an ideal relative to $\mathfrak{R}_i$, i.e., $z_ix_i, x_iz_i \in \mathfrak{S}_i$ for all $x_i \in \mathfrak{R}_i$, $z_i \in \mathfrak{S}_i$. Since $x_jz_i = z_ix_j = 0$ for $x_j \in \mathfrak{R}_j$ $(j \neq i)$, we have xz_i, $z_ix \in \mathfrak{S}_i$, and so $\mathfrak{S}_i$ is an ideal of $\mathfrak{R}$ and hence an invariant subspace relative to $\mathfrak{A}$. This contradicts the irreducibility of $\mathfrak{R}_i$.

We recall that an algebra $\mathfrak{A}$ of linear transformations in a vector space $\mathfrak{R}$ is completely reducible if it is semi-simple. Suppose conversely that $\mathfrak{A}$ is completely reducible, say, $\mathfrak{R} = \mathfrak{R}_1 \oplus \cdots \oplus \mathfrak{R}_k$, where the $\mathfrak{R}_i$ are irreducible invariant subspaces, and let $\mathfrak{N}$ be a nilpotent ideal of $\mathfrak{A}$. If $\mathfrak{B}$ is any subalgebra of $\mathfrak{A}$, and $\mathfrak{S}$ a subspace of $\mathfrak{R}$, we denote the subspace of elements ΣyB, $y \in \mathfrak{S}$, $B \in \mathfrak{B}$ by $\mathfrak{S}\mathfrak{B}$. Since $(\mathfrak{S}\mathfrak{B})\mathfrak{A} = \mathfrak{S}(\mathfrak{B}\mathfrak{A})$, $\mathfrak{S}\mathfrak{B}$ is invariant if $\mathfrak{B}$ is a right ideal. In particular $\mathfrak{R}_i\mathfrak{N}$ is invariant, and since the $\mathfrak{R}_i$ are irreducible, either $\mathfrak{R}_i\mathfrak{N} = 0$ or $\mathfrak{R}_i\mathfrak{N} = \mathfrak{R}_i$. But $\mathfrak{R}_i\mathfrak{N} = \mathfrak{R}_i$ implies $\mathfrak{R}_i = \mathfrak{R}_i\mathfrak{N}^\rho = 0$ if ρ is sufficiently high. Thus $\mathfrak{R}_i\mathfrak{N} = 0$ and $\mathfrak{R}\mathfrak{N} = 0$, i.e., $\mathfrak{N} = 0$ and so $\mathfrak{A}$ is semi-simple. By Theorem 1 we have therefore

THEOREM 2. *A necessary and sufficient condition that $\mathfrak{R}$ be a direct sum of simple algebras is that $\mathfrak{A}$ be semi-simple.*

[3] For definitions of *irreducibility, direct sum, complete reducibility, equivalence* (operator-isomorphism) for systems of linear transformations, see van der Waerden's *Moderne Algebra*, vol. II, 1931, §108. We shall also require a number of results on the structure and representation of semi-simple algebras. These may be found in §§115, 116, 118, 119, 121 of van der Waerden's book.

An element z is an *absolute zero-divisor* if $z \neq 0$ and $zx = 0 = xz$ for all x in $\mathfrak{R}$. Consider $\mathfrak{R}_\Sigma$ where Σ is an extension of Φ and suppose that $z' = a_1\zeta_1' + \cdots + a_n\zeta_n'$ is an absolute zero-divisor. If $a_ia_j = \Sigma a_k\gamma_{kij}$ ($\gamma \in \Phi$), then $z'a_i = a_iz' = 0$ implies that $\sum_i \gamma_{kij}\zeta_i' = 0$ and $\sum_j \gamma_{kij}\zeta_j' = 0$. Since these linear homogeneous equations have a non-trivial solution $\zeta_1', \cdots, \zeta_n'$ in Σ, they also have one, say, $\zeta_1, \cdots, \zeta_n$ in Φ, and so $\Sigma a_i\zeta_i$ is an absolute zero-divisor in $\mathfrak{R}$. Thus $\mathfrak{R}_\Sigma$ has absolute zero-divisors if and only if $\mathfrak{R}$ has. We note also that a simple algebra $\mathfrak{R}$ has no absolute zero-divisors unless $\mathfrak{R} = z\Phi$ where $z^2 = 0$. We suppose from now on that $\mathfrak{R}$ has no absolute zero-divisors. With this restriction we have

THEOREM 3. *If $\mathfrak{R} = \mathfrak{R}_1 \oplus \cdots \oplus \mathfrak{R}_k$ and the $\mathfrak{R}_i$ are simple, then $\mathfrak{A} = \mathfrak{A}_1 \oplus \cdots \oplus \mathfrak{A}_k$ and the $\mathfrak{A}_i$ are simple, and conversely. $\mathfrak{A}_i$ is the enveloping algebra of the left and right multiplications of $\mathfrak{R}_i$ (acting in $\mathfrak{R}$).*

Let $\mathfrak{R} = \mathfrak{R}_1 \oplus \cdots \oplus \mathfrak{R}_k$ and $\mathfrak{A}_i$ be the enveloping algebra of the multiplications of $\mathfrak{R}_i$. The elements of $\mathfrak{A}_i$ map $\mathfrak{R}_j$ ($j \neq i$) on 0 and $\mathfrak{R}$ on a subspace $\neq 0$ of $\mathfrak{R}_i$. It follows directly that $\mathfrak{A} = \mathfrak{A}_1 \oplus \cdots \oplus \mathfrak{A}_k$. $\mathfrak{A}_i \neq 0$ since the elements of $\mathfrak{R}_i$ are not absolute zero-divisors. Since the transformations of $\mathfrak{A}$ map $\mathfrak{R}_j$ on 0, the algebra $\mathfrak{A}_i$ is isomorphic to the enveloping algebra of the multiplications of $\mathfrak{R}_i$ acting in $\mathfrak{R}_i$. The latter is simple since it is an irreducible system of linear transformations and hence $\mathfrak{A}_i$ is simple also. Conversely if $\mathfrak{A} = \mathfrak{A}_1 \oplus \cdots \oplus \mathfrak{A}_k$, where the $\mathfrak{A}_i$ are simple, $\mathfrak{A}$ is completely reducible and hence $\mathfrak{R} = \mathfrak{R}_1 \oplus \cdots \oplus \mathfrak{R}_{k'}$, where the $\mathfrak{R}_i$ are simple algebras. By the first part and the uniqueness of the decomposition of an algebra as a direct sum of simple algebras we conclude that $k = k'$ and $\mathfrak{A}_i$ is the enveloping algebra of the multiplications of $\mathfrak{R}_i$.

COROLLARY. *$\mathfrak{R}$ is simple if and only if $\mathfrak{A}$ is.*

2. Let $\mathfrak{C}$ denote the centrum of $\mathfrak{A}$. If $\mathfrak{R}$ is itself associative and has an identity, $\mathfrak{C}$ coincides with the multiplications determined by the elements of the centrum $\mathfrak{C}'$ of $\mathfrak{R}$. For if $c \in \mathfrak{C}'$, $C_r = C_l \in \mathfrak{C}$, and if $C \in \mathfrak{C}$ and $1C = c$, then $xC = (1x)C = (1C)x = cx = (x1)C = x(1C) = xc$, so that $c \in \mathfrak{C}'$ and $C = C_r = C_l$. If $\mathfrak{R}$ is associative but has no identity, we may adjoin an identity to it and repeat the argument. We then obtain the fact that $\mathfrak{C}$ is the algebra of linear transformations determined by the multiplications of $\mathfrak{C}'$ plus the identity mapping. When $\mathfrak{R}$ is arbitrary we shall call $\mathfrak{C}$ the *extended centrum* of $\mathfrak{R}$.

If $\mathfrak{R}$ is simple, so is $\mathfrak{A}$, and hence $\mathfrak{C} = \mathrm{P}$ is an algebraic field of finite order over Φ. If $\xi \in \mathrm{P}$,

$$(xy)\xi = (x\xi)y = x(y\xi),$$

and so $\mathfrak{R}$ may be regarded as an algebra over P.

An algebra $\mathfrak{R}$ over Φ will be called *normal simple* if $\mathfrak{R}_\Omega$, the algebra obtained by extending Φ to its algebraic closure Ω, is simple.

THEOREM 4. *$\mathfrak{R}$ is normal simple if and only if it is simple and its extended centrum consists of the multiples of the identity transformation.*

By hypothesis the centrum of the simple algebra $\mathfrak{A}$ consists of the Φ-multiples of 1. It is a well-known result that $\mathfrak{A}_\Omega$ is simple in this case. But $\mathfrak{A}_\Omega$ is the enveloping algebra of the multiplications of $\mathfrak{R}_\Omega$, and hence by the corollary to Theorem 3 the latter is simple. On the other hand, if $\mathfrak{C}$ is larger than Φ, $\mathfrak{A}_\Omega$ is not simple[4] and hence $\mathfrak{R}$ is not normal simple.

Thus if $\mathfrak{R}$ is an arbitrary simple algebra, it becomes normal simple when regarded as an algebra over its extended centrum.

THEOREM 5. *If $\mathfrak{R}$ is simple and has order n over its extended centrum* P, *then $\mathfrak{A} \cong \mathrm{P}_n$, the algebra of n-rowed square matrices with coefficients in* P.

We regard P as the underlying field. Since $\mathfrak{R}_\Omega$ is simple, $\mathfrak{A}$ is an absolutely irreducible system of linear transformations. It follows from Burnside's theorem that $\mathfrak{A}$ contains n^2 linearly independent linear transformations and hence is isomorphic to P_n.

More generally the structure of $\mathfrak{A}$ when $\mathfrak{R}$ is a direct sum of simple algebras may be deduced from Theorems 3 and 5.

3. Now suppose that $\mathfrak{R}$ is an associative algebra with an identity. It is well known that the right multiplications form an algebra $\mathfrak{R}_r$ isomorphic to $\mathfrak{R}$ and the left multiplications form an algebra $\mathfrak{R}_l$ reciprocally isomorphic to $\mathfrak{R}$. $\mathfrak{R}_l$ ($\mathfrak{R}_r$) is the totality of linear transformations in the vector space $\mathfrak{R}$ commutative with those of $\mathfrak{R}_r$ ($\mathfrak{R}_l$). Thus $\mathfrak{R}_r \cap \mathfrak{R}_l = \mathfrak{C}$.

If $\mathfrak{R}$ is normal simple, $\mathfrak{R}_r \cap \mathfrak{R}_l = 1\Phi$. If the order of $\mathfrak{R}$ over Φ, $(\mathfrak{R}:\Phi) = n$ by Theorem 5, $(\mathfrak{A}:\Phi) = n^2 = (\mathfrak{R}_r:\Phi)(\mathfrak{R}_l:\Phi)$. Thus $\mathfrak{A}$ is a direct product of $\mathfrak{R}_r$ and $\mathfrak{R}_l$ and so we have obtained an elementary proof of the following theorem due to Brauer:[5]

THEOREM 6. *The direct product of a normal simple algebra and its reciprocal algebra is a complete matric algebra.*

4. We return to the general case in which $\mathfrak{R}$ is not necessarily associative and suppose that $x \to x^S$ is an automorphism of $\mathfrak{R}$ over Φ, i.e.,

$$(x + y)^S = x^S + y^S, \qquad (x\alpha)^S = x^S\alpha, \qquad (xy)^S = x^S y^S,$$

and $x \to x^S$ is (1-1). If $P = \sum A_{1i_1} \cdots A_{si_s}$ ($i_a = r$ or l) is an element of $\mathfrak{A}$, we define $P^S = \sum A^S_{1i_1} \cdots A^S_{si_s}$, where A^S_i is the right or left multiplication determined by a^S. P^S is independent of the representation of P. For if $\sum A_{1i_1} \cdots A_{si_s} = \sum B_{1j_1} \cdots B_{tj_t}$ ($j_\beta = r, l$), i.e.,

$$x(\sum A_{1i_1} \cdots A_{si_s}) = x(\sum B_{1j_1} \cdots B_{tj_t})$$

[4] If $\mathfrak{C}$ is separable, $\mathfrak{A}_\Omega$ is semi-simple though not simple, and if $\mathfrak{C}$ is inseparable, $\mathfrak{A}_\Omega$ has a radical. Cf. van der Waerden, loc. cit., §119.

[5] R. Brauer, *Über Systeme hyperkomplexer Zahlen*, Math. Zeits., vol. 30 (1929), p. 103.

for all x, then

$$x^S(\sum A^S_{1i_1} \cdots A^S_{si_s}) = x^S(\sum B^S_{1j_1} \cdots B^S_{tj_t})$$

for all x^S. Since x^S ranges over all of $\mathfrak{R}$ when x does, we have $\sum A^S_{1i_1} \cdots A^S_{si_s} = \sum B^S_{1j_1} \cdots B^S_{tj_t}$. It follows that the correspondence $P \rightarrow P^S$ is an automorphism of $\mathfrak{A}$ and $(xP)^S = x^S P^S$.

Any automorphism of an associative algebra induces an automorphism in its centrum. Hence if $\mathfrak{R}$ is simple, an automorphism of $\mathfrak{R}$ defines an automorphism in the field $\mathfrak{C} = \mathrm{P}$, the extended centrum.

THEOREM 7. *If $\mathfrak{R}$ is normal simple over P and $\mathrm{P} \supset \Phi$ such that $(\mathrm{P}:\Phi)$ is finite, then the automorphisms of $\mathfrak{R}$ over Φ have the property $(x\xi)^S = x^S\xi^S$, where $\xi \in \mathrm{P}$ and $\xi \rightarrow \xi^S$ is an automorphism of P.*

Let $\mathfrak{G}$ be the group of automorphisms of $\mathfrak{R}$ over Φ and $\mathfrak{T}$ the subgroup consisting of the automorphisms of $\mathfrak{R}$ over P. Theorem 7 shows that $\mathfrak{T}$ is an invariant subgroup of $\mathfrak{G}$ and that $\mathfrak{G}/\mathfrak{T}$ is isomorphic to a subgroup of the Galois group $\mathfrak{g}$ of P over Φ.

Now suppose that $\mathfrak{R} = \mathfrak{R}_0 \times \mathrm{P}$ ($= \mathfrak{R}_{0\mathrm{P}}$ regarded as an algebra over Φ) where $\mathfrak{R}_0$ is a normal simple algebra over Φ. If $a_1, \cdots, a_n$ is a basis for $\mathfrak{R}_0$ over Φ or for $\mathfrak{R}$ over P and S is an element of $\mathfrak{g}$, then the correspondence $x = \sum a_i\xi_i \rightarrow \sum a_i\xi_i^S = x^{S_1}$ is an automorphism of $\mathfrak{R}$ such that $(x\xi)^{S_1} = x^{S_1}\xi^{S_1} = x^{S_1}\xi^S$. Let $\mathfrak{G}_1$ denote the subgroup of $\mathfrak{G}$ consisting of the elements S_1. Evidently $\mathfrak{G}_1 \cong \mathfrak{g}$ and $\mathfrak{G}_1 \cap \mathfrak{T} = I$ the identity mapping. By Theorem 7 any element of $\mathfrak{G}$ has the form S_1H where $S_1 \in \mathfrak{G}_1$ and $H \in \mathfrak{T}$. Hence $\mathfrak{G}/\mathfrak{T} \cong \mathfrak{G}_1 \cong \mathfrak{g}$. This result may be used, as we shall show in another paper, to determine the automorphisms of simple Lie algebras and simple continuous groups.

UNIVERSITY OF CHICAGO

THE INVERSION PROBLEM OF MÖBIUS

By Einar Hille

1. **Introduction.** The present paper represents an attempt to give a rigorous treatment of certain inversion problems which have their origin in a little-known paper by A. F. Möbius.[1]

As a typical, though not the oldest, example of these inversion problems we might take the linear functional equation with constant coefficients

$$\sum_{n=1}^{\infty} a_n f(nz) = g(z), \tag{1.1}$$

a formal solution of which has the form

$$\sum_{n=1}^{\infty} b_n g(nz) = f(z). \tag{1.2}$$

These problems all lead to the same infinite system of bilinear equations

$$a_1 b_1 = 1, \qquad \sum_{d \mid n} a_d b_{n/d} = 0, \quad n > 1, \tag{1.3}$$

for which the *algorithm of Möbius* seems a fitting name.

This algorithm is perhaps best known from the problem of finding the reciprocal of an ordinary Dirichlet series, i.e., a solution of the problem

$$\sum_{n=1}^{\infty} a_n n^{-s} \sum_{n=1}^{\infty} b_n n^{-s} = 1. \tag{1.4}$$

We shall see that the properties of these series are fundamental in all these inversion problems.

This observation suggests that there is a class of inversion problems associated with the problem of expressing the reciprocal of a general Dirichlet series or, still more generally, of a Laplace-Stieltjes integral as a function of the same class. In general the reciprocal is not so expressible, but whenever it is, certain functional equations of the type

$$\int_1^{\infty} f(uz)\, dA(u) = g(z) \tag{1.5}$$

have solutions of the form

$$\int_1^{\infty} g(uz)\, dB(u) = f(z), \tag{1.6}$$

Received April 24, 1937; presented to the American Mathematical Society, March 26, 1937.

[1] *Ueber eine besondere Art von Umkehrung der Reihen*, Journal f. Math., vol. 9 (1832), pp. 105–123; *Gesammelte Werke*, vol. IV, 1887, pp. 589–612.

where

$$\int_1^u A\left(\frac{u}{v}\right) dB(v) = 1, \quad 1 < u. \tag{1.7}$$

The last equation is the transcendental analogue of the algorithm of Möbius.

In §2 of the present paper there is a discussion of the original problem of Möbius, of the algorithm of Möbius and of the problem of finding the reciprocal of an ordinary Dirichlet series. While there is comparatively little that is strictly new in this paragraph, the results are necessary for the rest of the paper and do not appear to be well known. In §3 we discuss equation (1.1) and various connected problems. In §4 we discuss equation (1.5) and the problem of expressing the reciprocal of a Laplace-Stieltjes integral as an integral of the same kind.

2. Some classical problems.

2.1. *The algorithm of Möbius.* Let $\mathfrak{A} = \{a_n\}$ be a given infinite sequence of real or complex numbers. The sequence is *proper* or *improper* according as $a_1 \neq 0$ or $= 0$, and in the former case it is *normalized* if $a_1 = 1$. Following O. Hölder,[2] we call $\mathfrak{B} = \{b_n\}$ *the reciprocal sequence* of $\mathfrak{A}$ if the latter is proper and $\mathfrak{A}$ and $\mathfrak{B}$ satisfy the algorithm of Möbius

$$a_1 b_1 = 1, \qquad \sum_{d \mid n} a_d b_{n/d} = 0, \quad n > 1, \tag{2.1.1}$$

or symbolically $\mathfrak{A}\mathfrak{B} = 1$. The underlying product definition is that of *Dirichlet multiplication*, i.e., in general $\mathfrak{A}\mathfrak{B} = \mathfrak{C}$, where the sequence $\mathfrak{C} = \{c_n\}$ is defined by

$$c_n = \sum_{d \mid n} a_d b_{n/d}. \tag{2.1.2}$$

Thus, formally,

$$\sum_{n=1}^{\infty} a_n n^{-s} \cdot \sum_{n=1}^{\infty} b_n n^{-s} = \sum_{n=1}^{\infty} c_n n^{-s}.$$

In case of the reciprocal sequence, $\mathfrak{C}$ is simply the unit sequence $1, 0, 0, \cdots$.

The system (2.1.1) determines $\mathfrak{B}$ uniquely. We have

$$b_n = \sum (-1)^{\alpha_1 + \alpha_2 + \cdots} C_{\alpha_1 \alpha_2 \cdots} (a_{d_1})^{\alpha_1} (a_{d_2})^{\alpha_2} \cdots, \tag{2.1.3}$$

where the summation extends over all factorizations of $n = d_1^{\alpha_1} d_2^{\alpha_2} \cdots$, and $C_{\alpha_1 \alpha_2 \cdots}$ is the combinatorial function which gives the number of possible arrangements of a set consisting of α_1 objects of one kind, α_2 objects of a second, etc. We have

$$\sum (-1)^{\alpha_1 + \alpha_2 + \cdots} C_{\alpha_1 \alpha_2 \cdots} = \mu(n), \tag{2.1.4}$$

[2] *Über gewisse der Möbiusschen Funktion $\mu(n)$ verwandte zahlentheoretische Funktionen, die Dirichletsche Multiplikation und eine Verallgemeinerung der Umkehrungsformeln,* Berichte d. Sächs. Akad. d. Wiss., Math.-phys. Kl., vol. 85 (1933), pp. 1–28.

the Möbius' μ-function, whereas

$$\sum C_{\alpha_1 \alpha_2 \cdots} = \pi(n), \tag{2.1.5}$$

the number of factorizations of n into factors $\neq 1$ ($n \neq 1$, $\pi(1) = 1$) when attention is paid to the order of the factors. $\pi(n)$ is a highly irregular function.[3] For the following it is enough to note that

$$\sum_{n=1}^{\infty} \pi(n) n^{-s} = [2 - \zeta(s)]^{-1} \tag{2.1.6}$$

for $\Re(s) > \rho$, $\zeta(\rho) = 2$, and that

$$\pi(n) < C_1 n^{\rho} \tag{2.1.7}$$

for all values of n, whereas for every $\epsilon > 0$ there are infinitely many values of n for which

$$\pi(n) > C_2 n^{\rho - \epsilon}. \tag{2.1.8}$$

2.2. *The reciprocation problem for ordinary Dirichlet series.* That the reciprocal of an ordinary Dirichlet series with $a_1 \neq 0$ can be represented by a series of the same kind is well known. The best theorem in this connection is one due to E. Landau.[4]

THEOREM 2.2.1. *Let*

$$D(s; a_n) = \sum_{n=1}^{\infty} a_n n^{-s}, \quad a_1 \neq 0 \tag{2.2.1}$$

have a domain of convergence, and let the function represented by the series be holomorphic and different from zero for $\sigma > \alpha$. Then

$$[D(s; a_n)]^{-1} \equiv \sum_{n=1}^{\infty} b_n n^{-s} \tag{2.2.2}$$

is convergent for $\sigma > \alpha$.

This theorem lies quite deep. It naturally brings up the question whether it is possible to assign upper bounds for the real parts of the possible zeros of $D(s; a_n)$ and thus also for the abscissa of convergence of the reciprocal. The answer is in the affirmative and is fairly trivial.

THEOREM 2.2.2. *Let $\{r_n\}$ be a given sequence of positive numbers, $r_1 = 1$, $r_n = O(n^{\kappa})$ for some fixed real κ. Consider the class $\mathfrak{D}$ of all Dirichlet series $D(s; a_n)$ with $a_1 = 1$, $|a_n| = r_n$, $n \geqq 2$. Let S be the abscissa of convergence of $D(s; r_n)$ and put $D(S + 0; r_n) = R \leqq \infty$. If $R > 2$, the equation*

$$D(\sigma; r_n) = 2 \tag{2.2.3}$$

[3] See E. Hille, *A problem in "factorisatio numerorum"*, Acta Arithmetica, vol. 2 (1936), pp. 134–144. $\pi(n)$ is denoted by $f(n)$ in this paper.

[4] *Über den Wertevorrat von $\zeta(s)$ in der Halbebene $\sigma > 1$*, Göttinger Nachrichten, 1933, pp. 81–91, p. 90.

has a real root $\rho = \rho(\mathfrak{T}) > S$. *No series of* $\mathfrak{T}$ *has any zeros in the half-plane* $\sigma > \rho$, *whereas there exist series in* $\mathfrak{T}$ *having infinitely many zeros in the strip* $\rho - \epsilon < \sigma < \rho$ *for every* $\epsilon > 0$. *If, on the other hand,* $1 < R \leqq 2$, *there are no zeros of any series in* $\mathfrak{T}$ *for* $\sigma > S$ *and there are series having either zeros or singular points in every strip* $S - \epsilon < \sigma < S$.

Proof. The verification of the fact that the series in $\mathfrak{T}$ cannot have zeros in the half-planes $\sigma > \rho(\mathfrak{T})$ and $\sigma > S$, respectively, is elementary and may be left to the reader. Further, the series $[D(s; a_n)]^{-1} \equiv D(s; b_n)$ are easily shown to be absolutely convergent in the same half-planes.

If $R > 2$, we note that the series

$$1 - \sum_{n=1}^{\infty} r_n n^{-s}$$

is a member of the class $\mathfrak{T}$ and vanishes at $s = \rho(\mathfrak{T})$. It further has infinitely many zeros in any strip $\rho(\mathfrak{T}) - \epsilon < \sigma < \rho(\mathfrak{T})$. If $1 < R \leqq 2$, the point $s = S$ is a non-polar singularity of $D(s; r_n)$ and consequently also a singularity of $[D(s; r_n)]^{-1}$. It follows that in either case the estimates given are the best possible valid for the whole class $\mathfrak{T}$.

It would be of some interest to know if these estimates for the upper bound of the real parts of the zeros are imposed upon us by a relatively small set of elements in $\mathfrak{T}$ or if they represent the rule rather than the exception. A discussion of this question in general calls for an interpretation of $\mathfrak{T}$ as a topological space, i.e., a definition of closure, possibly based upon a definition of distance or of measure.

But there is one very special case in which a complete answer is available without any topology. Suppose that $a_n = 0$ unless n is a prime, and put

$$a_{p_k} = \alpha_k\,, \qquad r_{p_k} = \rho_k\,, \qquad D(s; a_n) = P(s; \alpha_k), \qquad \mathfrak{T} = \mathfrak{P}.$$

Let us suppose that $R > 2$. Using a classical theorem of H. Bohr on the relation between the set of values of a Dirichlet series and of the associated power series in infinitely many unknowns,[5] we conclude that every series $P(s; \alpha_k)$ has infinitely many zeros in every strip $\rho(\mathfrak{P}) - \epsilon < \sigma < \rho(\mathfrak{P})$. Indeed, the associated power series is simply the linear form

$$L(x) = 1 + \sum_{1}^{\infty} \alpha_k x_k = 1 + \sum_{1}^{\infty} \rho_k e^{i\theta_k} x_k\,,$$

and putting $x_k = -e^{-i\theta_k} p_k^{-\rho}$, we get $L(x) = 1 - \sum_{1}^{\infty} \rho_k p_k^{-\rho} = 0$. By Bohr's theorem the value 0 is taken on infinitely often by $P(s; \alpha_k)$ in every strip $\rho - \epsilon < \sigma < \rho$. Hence in this case all the reciprocal series in $\mathfrak{P}$ have the same abscissa of convergence, viz., $\rho(\mathfrak{P})$.

[5] *Über die Bedeutung der Potenzreihen unendlich vieler Variabeln in der Theorie der Dirichletschen Reihen* $\sum a_n n^{-s}$, Göttinger Nachrichten, 1913, pp. 441–488, p. 451.

2.3. *Möbius' problem.* Möbius[6] raised the following question: given a power series

$$f(z) = \sum_{n=1}^{\infty} a_n z^n, \quad a_1 \neq 0, \tag{2.3.1}$$

find the expansion of z in terms of the functions $f(z^n)$, $n = 1, 2, 3, \cdots$. Let it be

$$z = \sum_{n=1}^{\infty} b_n f(z^n). \tag{2.3.2}$$

A straightforward calculation shows that the b's must satisfy the algorithm of Möbius. Further, it is clear that if

$$F(z) = \sum_{n=1}^{\infty} A_n z^n, \tag{2.3.3}$$

then

$$F(z) = \sum_{n=1}^{\infty} B_n f(z^n), \tag{2.3.4}$$

where

$$B_n = \sum_{d \mid n} b_d A_{n/d}. \tag{2.3.5}$$

All this is highly formal and an analyst naturally wants to know the range of validity of the formulas, conditions for convergence, etc.

As a preliminary step in this study, let us suppose that the power series in (2.3.1) has a circle of convergence, and form the adjoint power series

$$\varphi(z) = \sum_{n=1}^{\infty} b_n z^n. \tag{2.3.6}$$

We call $\varphi(z)$ the Möbius transform of $f(z)$,

$$\varphi(z) = \mathfrak{M}[f(z)]. \tag{2.3.7}$$

The Möbius algorithm shows that conversely $f(z)$ is the Möbius transform of $\varphi(z)$,

$$\mathfrak{M}[\mathfrak{M}[f(z)]] = f(z), \tag{2.3.8}$$

i.e., the Möbius transformation is an involution.

We must show that the power series in (2.3.6) is also convergent. This is established in

THEOREM 2.3.1. *Let the radii of convergence of the power series in* (2.3.1) *and* (2.3.6) *be* R_1 *and* R_2 *respectively. If* $0 < R_1 < 1$, *then* $R_1 = R_2$; *if* $1 \leq R_1$, *then also* $1 \leq R_2$.

[6] Loc. cit., *Werke*, vol. IV, p. 591.

Proof. Choose $R_0 < R_1$. Then there exists an M such that $|a_n| \leqq MR_0^{-n}$ for every n. Hence by (2.1.3)

$$|b_n| \leqq \sum C_{\alpha_1\alpha_2}\dots M^{\alpha_1+\alpha_2+\cdots} R_0^{-(\alpha_1 d_1+\alpha_2 d_2+\cdots)}.$$

If $R_1 \leqq 1$, then $R_0 < 1$. We can assume $M \geqq 1$ without restricting the generality. Further,

$$\alpha_1 d_1 + \alpha_2 d_2 + \cdots \leqq n, \qquad \alpha_1 + \alpha_2 + \cdots \leqq p(n),$$

where $p(n)$ denotes the total number of prime factors of n. Hence by (2.1.5)

$$|b_n| \leqq \pi(n)M^{p(n)}R_0^{-n}. \tag{2.3.9}$$

Since $p(n) = O(\log n)$ and $\pi(n) = O(n^\rho)$, we conclude that $R_0 \leqq R_2$ or $R_1 \leqq R_2$.

Suppose next that $R_1 > 1$. We can then choose $R_0 > 1$. Further,

$$\alpha_1 d_1 + \alpha_2 d_2 + \cdots \geqq \nu_1 p_{i_1} + \nu_2 p_{i_2} + \cdots \equiv P(n)$$

if $n = p_{i_1}^{\nu_1} p_{i_2}^{\nu_2} \cdots$, where the p_{i_k} are the distinct prime factors of n. Hence

$$|b_n| \leqq \pi(n)M^{p(n)}R_0^{-P(n)}. \tag{2.3.10}$$

But $P(n)$ is infinitely often $o(n)$. It follows that

$$\overline{\lim_{n\to\infty}}\, |b_n|^{1/n} \leqq 1,$$

or $R_2 \geqq 1$.

In order to complete the proof for the case $R_1 < 1$, we use the involutory character of the transformation. We have shown that $R_1 \leqq R_2$. If $R_2 \leqq 1$, we can conclude that $R_2 \leqq R_1$, i.e., $R_1 = R_2$, simply by noticing that $f(z)$ is the Möbius transform of $\varphi(z)$. On the other hand, the assumption $R_2 > 1$ implies by the same argument that $R_1 \geqq 1$, and this contradicts the original assumption. Hence $R_1 < 1$ implies $R_1 = R_2$.

If $R_1 = 1$, we may well have $R_2 > 1$. The situation becomes clearer by introducing the associated Dirichlet series

$$D(s; a_n) = \sum_{n=1}^{\infty} a_n n^{-s}, \qquad D(s; b_n) = \sum_{n=1}^{\infty} b_n n^{-s}.$$

The assumption $R_2 > 1$ implies that $D(s; b_n)$ converges for all s and is an entire function of s. Hence $D(s; a_n)$ has a half-plane of absolute convergence and is not merely a formal Dirichlet series; moreover, $a_n = O(n^\kappa)$ for some finite value of κ.

Thus if $R_1 = 1$, we have $R_2 = R_1$ unless $D(s; b_n)$ is an entire function, and then $R_2 \geqq R_1$. This case can arise only when $a_n = O(n^\kappa)$.

If $R_1 > 1$, $D(s; a_n)$ is an entire function of s, and normally $R_2 = 1$, unless $D(s; a_n) \neq 0$, in which case $R_2 \geqq 1$.

After this discussion it is easy to discuss the validity of Möbius' inversion formula.

THEOREM 2.3.2. *If R_1 is the radius of convergence of the power series for $f(z)$,*

the inversion formula (2.3.2) *is valid for* $|z| < \min(R_1, 1)$. *If* $R_1 < 1$, *the series diverges for* $R_1 < |z| < 1$. *The series may converge outside of the unit circle, but normally it does not represent* z *for such values. If the radius of convergence of* $F(z)$ *in* (2.3.3) *is* R, *formula* (2.3.4) *is valid for* $|z| < \min(R, R_1, 1)$.

Proof. Let $0 < R_1 \leqq 1$. Then all the terms in (2.3.2) are regular analytic functions of z in $|z| < R_1$. By (2.3.9)

$$\begin{aligned}\sum_{n=1}^{\infty} |b_n f(z^n)| &\leqq \sum_{n=1}^{\infty} |b_n| \sum_{m=1}^{\infty} |a_m| \cdot |z|^{mn} \\ &\leqq \sum_{n=1}^{\infty} \pi(n) M^{p(n)+1} R_0^{-n} \frac{|z|^n}{1 - |z|^n/R_0},\end{aligned} \tag{2.3.11}$$

and this is clearly convergent for $|z| < R_0$. Here $R_0 < R_1$ and as near to R_1 as we please, i.e., the series in (2.3.2) is absolutely convergent for $|z| < R_1$. Moreover, the double series obtained by substituting the power series for $f(z^n)$ on the right side of (2.3.2) is absolutely convergent, as we have just seen. It can consequently be rearranged at liberty. Collecting powers of equal degree and reducing with the aid of the algorithm of Möbius, the double series reduces to its first term z. This completes the proof for the case $R_1 \leqq 1$.

If $R_1 > 1$, the terms of the series (2.3.2) are holomorphic for $|z| \leqq 1$ and in no larger region. For such values formula (2.3.10) shows that the double series is dominated by

$$\sum_{n=1}^{\infty} \pi(n) M^{p(n)+1} R_0^{-p(n)} \frac{|z|^n}{1 - |z|^n/R_0}. \tag{2.3.12}$$

It follows that the inversion formula is valid for $|z| < 1$.

Let $|z| < 1$. Then

$$\overline{\lim_{n\to\infty}} |b_n f(z^n)|^{1/n} = |z| \overline{\lim_{n\to\infty}} |b_n|^{1/n} = |z|/R_2.$$

It follows that if $R_1 < 1$, so that $R_2 = R_1$, the series (2.3.2) diverges in the annulus $R_1 < |z| < 1$.

Outside of the unit circle the situation may differ considerably in different cases. Thus if

$$f(z) = \frac{z}{1-z}, \quad \text{then} \quad z = \sum_{n=1}^{\infty} \mu(n) \frac{z^n}{1-z^n},$$

which clearly diverges for $|z| > 1$. But if

$$f(z) = \frac{z}{1-z^2},$$

then

$$b_n = \begin{cases} \mu(2k+1), & n = 2k+1, \\ 0, & n = 2k, \end{cases}$$

so that the series

$$\sum_{n=1}^{\infty} b_n \frac{z^n}{1 - z^{2n}}$$

converges also outside of the unit circle, but to $-1/z$ instead of to z. Finally, if

$$f(z) = ze^{-z^k},$$

where k is a positive integer, then the series

$$\sum_{n=1}^{\infty} b_n z^n e^{-z^{nk}}$$

converges on the rays $|z| > 1$, $\arg z = \nu 2\pi/k$, $\nu = 0, 1, \cdots, k - 1$, and nowhere else outside of the unit circle. The sum of the series tends to zero as $z \to \infty$ along the rays in question; thus, the sum of the series cannot be z, but I am unable to determine its actual value.

The reader will have no difficulty in verifying the statements concerning $F(z)$ in Theorem 2.3.2 on the basis of the estimates for the coefficients, and this part of the proof will be omitted.

3. A class of linear functional equations.

3.1. *The Möbius $\mathfrak{A}$-transform.* Let $\mathfrak{A} = \{a_n\}$ be a given proper, normalized sequence, $\mathfrak{B} = \{b_n\}$ the reciprocal sequence in the sense of §2.1. Möbius[7] observed that the same algorithm enters in the study of the functional equation

$$G(z) = \sum_{n=1}^{\infty} a_n F(z^n) \tag{3.1.1}$$

for which he proposed the solution

$$F(z) = \sum_{n=1}^{\infty} b_n G(z^n). \tag{3.1.2}$$

Conversely, (3.1.1) is a solution of (3.1.2) if $F(z)$ is the given function. Though Möbius claims that these relations hold for arbitrary given functions, he has presumably only had power series in mind, and there is no indication that he looked into convergence questions at all.

Putting

$$G(e^z) = g(z), \qquad F(e^z) = f(z),$$

we can rewrite the functional equations as follows:

$$g(z) = \sum_{n=1}^{\infty} a_n f(nz), \tag{3.1.3}$$

$$f(z) = \sum_{n=1}^{\infty} b_n g(nz). \tag{3.1.4}$$

[7] Loc. cit., *Werke*, vol. IV, p. 593.

These forms are more convenient to handle than the original ones, and will serve as the basis of the discussion in the present paragraph. Special cases have long been in the literature. Thus the case $a_n = 1$, $b_n = \mu(n)$, $z = k$, gives the inversion formulas

$$g(k) = \sum_{n=1}^{\infty} f(nk), \tag{3.1.5}$$

$$f(k) = \sum_{n=1}^{\infty} \mu(n) g(nk), \tag{3.1.6}$$

which are used extensively in analytical number theory.[8]

We proceed to an analytical discussion of equation (3.1.3). Let E be a set of points in the complex plane such that if E contains the point z_0 it also contains all multiples nz_0 of z_0, $n = 2, 3, \cdots$. Let $f(z)$ be given in E and such that

$$\mathfrak{M}[f(z); \mathfrak{A}] \equiv \sum_{n=1}^{\infty} a_n f(nz) \tag{3.1.7}$$

converges in E. We call this function the Möbius $\mathfrak{A}$-transform of $f(z)$ with similar notation and terminology for $\mathfrak{B}$ and for other sequences. The reciprocity of the two sequences $\mathfrak{A}$ and $\mathfrak{B}$ is reflected in the property

$$\mathfrak{M}[\mathfrak{M}[f(z); \mathfrak{A}]; \mathfrak{B}] = \mathfrak{M}[\mathfrak{M}[f(z); \mathfrak{B}]; \mathfrak{A}] = f(z), \tag{3.1.8}$$

valid for sufficiently restricted classes of functions $f(z)$.

THEOREM 3.1.1. *A sufficient condition for the validity of* (3.1.8) *is the absolute convergence of the series*

$$S[f] \equiv \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} a_m b_n f(mnz). \tag{3.1.9}$$

Proof. The series, being absolutely convergent, can be rearranged arbitrarily. If summed by columns its sum is $\mathfrak{M}[\mathfrak{M}[f(z); \mathfrak{A}]; \mathfrak{B}]$, if summed by rows, $\mathfrak{M}[\mathfrak{M}[f(z); \mathfrak{B}]; \mathfrak{A}]$, whereas summation over constant values of the product mn gives simply $f(z)$, by virtue of Möbius' algorithm.

3.2. *Inversion of the $\mathfrak{A}$-transform.* We now turn to the question of finding the inverse of the $\mathfrak{A}$-transform, i.e., the resolution of the equation

$$\mathfrak{M}[f(z); \mathfrak{A}] = g(z) \tag{3.2.1}$$

for $f(z)$ in terms of $g(z)$.

THEOREM 3.2.1. *A sufficient condition that*

$$f(z) = \mathfrak{M}[g(z); \mathfrak{B}] \tag{3.2.2}$$

[8] See, e.g., P. Bachmann, *Die analytische Zahlentheorie*, Leipzig, 1894, p. 310 et seq. Bachmann does not seem to have been aware of Möbius' paper. Thus he credits the introduction of the function $\mu(n)$ to F. Mertens, *Ueber einige asymptotische Gesetze der Zahlentheorie*, Journal f. Math., vol. 77 (1874), pp. 289–338.

be a solution of (3.2.1) *which is absolutely convergent in E is that the series* $S[g]$ *be absolutely convergent in E. On the other hand, there can be at most one solution of* (3.2.1) *which renders the series* $S[f]$ *absolutely convergent, and whenever it exists this solution is given by* (3.2.2).

Proof. The assumption that $S[g]$ is absolutely convergent implies that $\mathfrak{M}[\mathfrak{M}[g(z); \mathfrak{B}]; \mathfrak{A}] = g(z)$ by Theorem 3.1.1. Hence (3.2.2) gives a solution under these circumstances, and the solution is evidently absolutely convergent. Conversely, if $f(z)$ is a solution of (3.2.1) such that $S[f]$ is absolutely convergent, then for the same reason

$$S[f] \equiv f(z) = \mathfrak{M}[\mathfrak{M}[f(z); \mathfrak{A}]; \mathfrak{B}] = \mathfrak{M}[g(z); \mathfrak{B}],$$

so the solution in question is uniquely determined and given by (3.2.2).

It must be granted that Theorem 3.2.1 is of a rather restrictive character. It should be pointed out, however, that the mere existence of $\mathfrak{M}[f(z); \mathfrak{A}]$ is not enough to insure that this function is a solution of (3.2.1). Thus, for example, if there exists a non-vanishing function $g(z)$ such that $\mathfrak{M}[g(z); \mathfrak{B}] \equiv 0$, then formula (3.2.2) certainly does not give a solution of (3.2.1). See further §3.3.

The following theorem is of a somewhat different character.

THEOREM 3.2.2. *Let* $g(z)$ *be holomorphic in the sector* S, $\theta_1 < \arg z < \theta_2$, $|z| > R > 0$, *and let*

$$g(z) = z^{-\alpha}\left[c_0 + O\left(\frac{1}{|z|}\right)\right] \quad \text{as } z \to \infty \text{ in } S. \tag{3.2.3}$$

Further, suppose that $D(s; a_n) \equiv \sum_1^\infty a_n n^{-s}$ *is convergent and different from zero for* $\Re(s) > \Re(\alpha) - \epsilon$, $\epsilon > 0$. *Then* (3.2.2) *defines a solution of* (3.2.1), *holomorphic in* S, *and*

$$f(z) = z^{-\alpha}\left[c_0[D(\alpha; a_n)]^{-1} + O\left(\frac{1}{|z|}\right)\right] \quad \text{as } z \to \infty \text{ in } S, \tag{3.2.4}$$

and this is the only solution of such asymptotic character.

Proof. We have to show that $\mathfrak{M}[\mathfrak{M}[g(z); \mathfrak{B}]; \mathfrak{A}] = g(z)$ under the given assumptions. We start by observing that

$$\mathfrak{M}[z^{-\alpha}; \mathfrak{A}] = D(\alpha; a_n)z^{-\alpha}, \tag{3.2.5}$$

$$\mathfrak{M}[z^{-\alpha}; \mathfrak{B}] = D(\alpha; b_n)z^{-\alpha}, \tag{3.2.6}$$

the convergence of $D(\alpha; b_n)$ being a consequence of Landau's Theorem 2.2.1. Let us put

$$g(z) = c_0 z^{-\alpha} + g_1(z), \tag{3.2.7}$$

$$f(z) = c_0 D(\alpha; b_n)z^{-\alpha} + f_1(z). \tag{3.2.8}$$

Then if $f(z)$ is a solution of (3.2.1), $f_1(z)$ is a solution of

$$\mathfrak{M}[f_1(z); \mathfrak{A}] = g_1(z). \tag{3.2.9}$$

But to this equation we can apply Theorem 3.2.1. Indeed, by assumption $g_1(z) = O(|z|^{-1-\gamma})$, $\gamma = \Re(\alpha)$, and the series $D(s; a_n)$ and $D(s; b_n)$ are absolutely convergent for $s = 1 + \gamma$, since they converge for $s = \gamma - \epsilon/2$. Forming $S[g_1]$ and replacing each term by its absolute value, we find that the series is dominated by a constant multiple of

$$|z|^{-1-\gamma} \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} |a_m b_n| (mn)^{-1-\gamma}$$

convergent in S. Hence $f_1(z) = \mathfrak{M}[g_1(z); \mathfrak{B}]$ is a solution of (3.2.9), and $\mathfrak{M}[g(z); \mathfrak{B}]$ is a solution of (3.2.1).

Further,

$$f_1(z) = O\left\{|z|^{-1-\gamma} \sum_{n=1}^{\infty} |b_n| n^{-1-\gamma}\right\},$$

or

$$|f_1(z)| \leqq M |z|^{-1-\gamma}, \tag{3.2.10}$$

whence it follows that the double series $S[f_1]$ is absolutely convergent in S. Hence, by Theorem 3.2.1, $f_1(z)$ is the only solution of (3.2.9) having this property; and, a fortiori, the only solution satisfying (3.2.10). It follows that $\mathfrak{M}[g(z); \mathfrak{B}]$ is the only solution of (3.2.1) of the form (3.2.8), where $f_1(z)$ satisfies (3.2.10).

It is obvious that the solution breaks down if $D(\alpha; a_n) = 0$. In general, it also breaks down if $D(s; a_n) = 0$ for an s with $\Re(s) > \Re(\alpha)$. It should be noted, however, that

$$[D(\alpha; a_n)]^{-1} z^{-\alpha}$$

is a solution of

$$\mathfrak{M}[f(z); \mathfrak{A}] = z^{-\alpha}$$

under the sole assumption that $D(\alpha; a_n)$ is convergent and different from zero. This observation may sometimes be used in order to extend the validity of our solution.

Thus, for example, Theorem 3.2.2 does not apply to the case $a_n = (-1)^{n-1}$ if $0 < \alpha < \frac{1}{2}$, but formula (3.2.2), nevertheless, gives a solution of the corresponding equation. This is readily seen by running over the proof again with this particular choice of the parameters.

Further, it should be noted that the assumption on the remainder in (3.2.3) is chosen merely with the view of insuring that the Dirichlet series $D(s; a_n)$ and $D(s; b_n)$ be absolutely convergent for $s = 1 + \gamma$. If there should exist a δ, $0 < \delta < 1$, such that these series converge absolutely for $s = \delta + \gamma$, it is sufficient for our purposes to assume that $z^{\alpha} g(z) = c_0 + O(|z|^{-\delta})$.

3.3. *Additional remarks.* We are dealing with two adjoint equations

$$\mathfrak{M}[x(z), \mathfrak{A}] = g(z), \tag{3.3.1}$$

$$\mathfrak{M}[y(z), \mathfrak{B}] = h(z), \tag{3.3.2}$$

and the corresponding homogeneous equations

$$\mathfrak{M}[u(z), \mathfrak{A}] = 0, \tag{3.3.3}$$

$$\mathfrak{M}[v(z), \mathfrak{B}] = 0. \tag{3.3.4}$$

We have already observed that

$$\mathfrak{M}[z^{-\alpha}, \mathfrak{A}] = D(\alpha; a_n)z^{-\alpha}, \tag{3.3.5}$$

provided $\mathfrak{R}(\alpha) > \sigma_0$, the abscissa of convergence of the series for $D(s; a_n)$. In the same domain we have

$$\mathfrak{M}\left[\frac{\partial^n}{\partial\alpha^n} z^{-\alpha}, \mathfrak{A}\right] = \frac{\partial^n}{\partial\alpha^n}[D(\alpha; a_n)z^{-\alpha}]. \tag{3.3.6}$$

From this we conclude that if the equation

$$D(s; a_n) = 0$$

has a k-fold zero at $s = \alpha$ with $\mathfrak{R}(\alpha) > \sigma_0$, then

$$z^{-\alpha}, z^{-\alpha} \log z, \cdots, z^{-\alpha} (\log z)^{k-1} \tag{3.3.7}$$

are solutions of the homogeneous equation (3.3.3).

It is clear that, if there are infinitely many zeros of $D(s; a_n)$ in the domain of convergence of the Dirichlet series, then any function of the form

$$\sum_{n=1}^{\infty} c_n z^{-\alpha_n} \tag{3.3.8}$$

satisfies (3.3.3), provided the double series

$$\sum_{m=1}^{\infty} a_m \sum_{n=1}^{\infty} c_n(mz)^{-\alpha_n}$$

can be rearranged so as to interchange the order of the summations.

It is perhaps worth while remarking that the series (3.3.8) do not form a dense set in any of the function spaces usually considered, such as $C[1, \infty]$ or $L_p(1, \infty)$, $1 \leqq p < \infty$. Indeed, the set $\{z^{-\alpha_n}\}$ will be closed in the space in question only if the series

$$\sum \frac{a + b\mathfrak{R}(\alpha_n)}{1 + |\alpha_n|^2}$$

diverges, where a and b are constants depending upon the space. But in our case $\mathfrak{R}(\alpha_n)$ is bounded; hence we are demanding the divergence of the series $\sum |\alpha_n|^{-2}$. But according to Landau[9] the number of zeros of an ordinary Dirichlet series in the domain $\sigma \geqq \sigma_0 + \epsilon$, $|t| \leqq T$, is $O(T \log T)$ and this frequency clearly does not permit the divergence of $\sum |\alpha_n|^{-2}$.

[9] E. Landau, *Über die Nullstellen der Dirichletschen Reihen*, Berliner Sitzungsberichte, vol. 14 (1913), pp. 897–907.

Our next remark concerns the existence of a solution of (3.3.1) when $g(z)$ satisfies (3.3.4). It is clear that the inversion formula of Möbius cannot give a solution in this case. It seems very plausible that no solution can exist under these circumstances. In certain simple cases it is possible to verify this surmise. Take, for example,

$$\sum_{k=0}^{\infty} f(2^k z) = 1.$$

Here $D(s; b_n) = 1 - 2^{-s}$ and $\mathfrak{M}[1, \mathfrak{B}] = 0$. Consider any domain E of the type described in §3.1, i.e., if $z_0 \epsilon E$, so do nz_0 for $n = 2, 3, \cdots$. If the equation holds for $z = z_0$ and for $z = 2z_0$, then we get by subtraction $f(z_0) = 0$. If it is true for $z = 2^k z_0$ for every integer k, we get by the same argument that $f(2^k z_0) = 0$ for every k. But this clearly contradicts the assumption that the equation holds for $z = z_0$. Thus the equation in question cannot have any solution in E or even in a point set S which contains $2z_0$ whenever it contains z_0.

The final remark of this paragraph concerns the solution of (3.3.1) when $g(z)$ is a solution of (3.3.3). It is enough to consider the case $g(z) = z^{-\alpha}$, where $s = \alpha$ is a k-fold root of the equation $D(s; a_n) = 0$ in the half-plane of convergence of the Dirichlet series. Equation (3.3.6) then shows that

$$x(z) = z^{-\alpha}\{(-1)^k [D^{(k)}(\alpha; a_n)]^{-1} (\log z)^k + \sum_{\nu=0}^{k-1} c_\nu (\log z)^\nu\}$$

is a solution of (3.3.1). This result shows a further analogy between the formal theory of the equations here considered and that of linear differential equations.

3.4. *Further equations with the same algorithm.* It was known to Möbius[10] that his algorithm entered in the study of other functional equations. The following example is slightly more general than the situation which Möbius had in mind.

We consider two multiplicative systems, i.e., we give two sequences $\{\epsilon(p)\}$ and $\{\omega(p)\}$, where p runs through the primes, we take $\epsilon(1) = \omega(1) = 1$ and define $\epsilon(n)$ and $\omega(n)$ by the equations

$$\epsilon(mn) = \epsilon(m)\epsilon(n), \qquad \omega(mn) = \omega(m)\omega(n). \tag{3.4.1}$$

Let us form the functional equation

$$\sum_{n=1}^{\infty} a_n \epsilon(n) f(\omega(n)z) = g(z). \tag{3.4.2}$$

It is not difficult to see that a formal solution is given by

$$f(z) = \sum_{n=1}^{\infty} b_n \epsilon(n) g(\omega(n)z), \tag{3.4.3}$$

[10] Möbius, loc. cit., *Werke*, vol. IV, p. 594.

where $\{b_n\}$ as usual is the reciprocal sequence of $\{a_n\}$. The elementary methods of §3.2 can be used to develop sufficient conditions for the validity of this inversion formula.[11] The details can be left to the reader.

4. General algorithms.

4.1. *The Möbius $A(u)$-transform.* We can write the Möbius $\mathfrak{A}$-transform as a Stieltjes integral, viz.,

$$\mathfrak{M}[f(z), \mathfrak{A}] = \int_1^\infty f(uz)\, dA(u),$$

where

$$A(u) = \sum_{n<u} a_n .$$

This suggests a generalization of the inversion problem to arbitrary functions $A(u)$ of bounded variation.

Let $A(u)$ be given for $u \geqq 1$, $A(1) = 0$, of bounded variation in every finite interval, and such that

$$A(u) = O(u^{\omega+\epsilon}) \tag{4.1.1}$$

for every $\epsilon > 0$, where $\omega \geqq 0$ is a constant.

Suppose now that $f(z)$ is an analytic function satisfying the following conditions: (i) $f(z)$ is holomorphic in a sectorial domain S such that if z is in S, so is uz for every $u \geqq 1$; (ii) the integral

$$\mathfrak{M}[f(z), A(u)] \equiv \int_1^\infty f(uz)\, dA(u) \tag{4.1.2}$$

exists in some domain $S_0 \subset S$.

We shall as a rule use the abbreviated notation $\mathfrak{M}[f, A]$ and refer to this function as the Möbius $A(u)$-transform of $f(z)$. To every fixed function $A(u)$ satisfying the above conditions there is a class $\mathfrak{F}[A]$ of functions $f(z)$ which admit $A(u)$-transforms in the sense of the definition.

The effective determination of $\mathfrak{F}[A]$ may be quite laborious except in the simplest cases, but it is easy to find a subclass of $\mathfrak{F}[A]$. Let $f(z)$ be holomorphic in a sectorial domain S and

$$|f'(z)| < M\,|z|^{-\gamma-1}, \qquad \gamma > \omega. \tag{4.1.3}$$

It is easy to see that $\mathfrak{M}[f, A]$ exists in S; thus every such function belongs to $\mathfrak{F}[A]$.

4.2. *The reciprocation problem for Laplace integrals.* In the case of the $\mathfrak{A}$-transform the inverse or reciprocal transform was given a priori and had a sense for a sufficiently restricted but not vacuous class of functions. For the

[11] Some instances of this inversion formula figured in the papers of E. Hille and O. Szász, *On the completeness of Lambert functions*, of which the first part appeared in the Bulletin of the American Mathematical Society, vol. 42 (1936), pp. 411–418, and the second in the Annals of Mathematics, vol. 37 (1936), pp. 800–815.

$A(u)$-transform the situation is different, and the inverse transform need not exist at all. By analogy with the sequence case we should consider the Laplace-Stieltjes integral

$$D(s) = \int_1^\infty u^{-s}\, dA(u). \tag{4.2.1}$$

The hypothesis (4.1.1) insures the convergence of this integral for $\sigma > \omega$. We should then take the reciprocal of $D(s)$ and find its representation as a Laplace-Stieltjes integral. But it is well known that $[D(s)]^{-1}$ ordinarily is not representable in this manner.

Necessary and sufficient conditions in order that $[D(s)]^{-1}$ shall be representable by a convergent Laplace-Stieltjes integral do not seem to be known. In the following I shall give two sets of two conditions each. One condition is common to the two sets; the first set is necessary, but perhaps not sufficient, the second set is sufficient, but certainly not necessary.[12]

THEOREM 4.2.1. *Let $D(s)$ be a function representable by a convergent Laplace-Stieltjes integral. In order that $[D(s)]^{-1}$ shall also admit such a representation, it is necessary that* (i) $\lim_{\sigma \to +\infty} D(\sigma) \neq 0$, *and* (ii) *there exist a half-plane $\sigma > \sigma_0$ in which $D(s) \neq 0$.*

Proof. That the conditions are necessary is obvious. Since

$$\lim_{\sigma \to +\infty} D(\sigma) = \lim_{u \to 1+} A(u),$$

we can replace condition (i) by the equivalent condition (i′) $A(u)$ is discontinuous at $u = 1$.

THEOREM 4.2.2. *If $A(u)$ is discontinuous at $u = 1$, $A(1 + 0) = a \neq 0$, and if the Laplace-Stieltjes integral representing $D(s)$ is absolutely convergent for $\sigma > \alpha$, then $[D(s)]^{-1}$ is representable by a Laplace-Stieltjes integral absolutely convergent for $\sigma >$* max (α, β), *where β is the root of the equation*

$$|a| = \int_1^\infty u^{-\sigma}\, dV_1^u[A(v) - a],^{13} \tag{4.2.3}$$

if it exists; otherwise $\beta = -\infty$.

Proof. Let us put

$$A_1(v) = A(v) - a, \qquad A_1(1) = 0,$$

$$V_1^u A_1(v) = A_0(u),$$

$$A_1(u, s) = \int_1^u v^{-s}\, dA_1(v),$$

$$A_0(u, s) = \int_1^u v^{-s}\, dA_0(v).$$

[12] It is to be hoped that the investigations by R. H. Cameron and N. Wiener, now in progress, will throw further light on this question.

[13] Here and in the following, $V_a^b f(t)$ denotes the total variation of $f(t)$ in $a \leq t \leq b$.

Then $A_0(u)$, $A_1(u)$, $A_0(u, s)$, and $A_1(u, s)$ are continuous at $u = 1$ and tend to zero as $u \to 1$. A simple consideration shows that

$$V_1^u A_1(v, s) \leqq A_0(u, \sigma). \tag{4.2.4}$$

By assumption

$$A_0(u) = O(u^{\alpha+\epsilon})$$

for every $\epsilon > 0$. It follows that for $\sigma > \alpha$ the increasing function $A_0(u, \sigma)$ tends to the finite limit $A_0(\infty, \sigma)$ as $u \to \infty$. Further, it is obvious that $A_0(\infty, \sigma)$ is monotone decreasing when σ increases and tends to zero as $\sigma \to \infty$. The latter conclusion follows from the fact that $A_0(u) \to 0$ monotonically from above as $u \to 0+$. Hence the equation

$$A_0(\infty, \sigma) = |\, a \,| \tag{4.2.5}$$

has at most one root $\geqq \alpha$. We define β to be equal to this root if it exists; otherwise we take $\beta = -\infty$.

Now let $\sigma > \max(\alpha, \beta)$. Then

$$D(s) = a + A_1(\infty, s),$$

and

$$|\, A_1(\infty, s) \,| \leqq A_0(\infty, \sigma) < |\, a \,|.$$

Hence

$$[D(s)]^{-1} = \sum_{n=0}^{\infty} (-1)^n a^{-n-1} [A_1(\infty, s)]^n, \tag{4.2.6}$$

and the series is absolutely convergent. We shall rewrite this series as a Laplace-Stieltjes integral. We have for $\sigma > \gamma > \max(\alpha, \beta)$

$$A_1(\infty, s) = \int_1^{\infty} u^{-(s-\gamma)} d_u A_1(u, \gamma). \tag{4.2.7}$$

Hence

$$[A_1(\infty, s)]^n = \int_1^{\infty} u^{-(s-\gamma)} d_u A_n(u, \gamma), \tag{4.2.8}$$

where

$$A_n(u, \gamma) = \int_1^u A_{n-1}(u/v, \gamma)\, d_v A_1(v, \gamma) \qquad (n = 2, 3, \cdots). \tag{4.2.9}$$

Here (4.2.8) is absolutely convergent, being the product of absolutely convergent Laplace-Stieltjes integrals.[14] This fact also follows from the subsequent estimates of $A_n(u, \gamma)$.

[14] For the properties of Laplace-Stieltjes integrals used in this paper, consult D. V. Widder, Trans. Amer. Math. Soc., vol. 31 (1929), pp. 694–743, and E. Hille and J. D. Tamarkin, Proc. Nat. Acad. Sci., vol. 19 (1933), pp. 573–577, 902–912; vol. 20 (1934), pp. 140–144.

We shall prove the inequality

$$(4.2.10) \qquad |A_n(u,\gamma)| \leqq V_1^u A_n(v,\gamma) \leqq [A_0(u,\gamma)]^n.$$

The inequality is obviously true for $n = 1$. Suppose that it has been proved for $n = k$. It follows in particular that $A_k(u, \gamma)$ is continuous at $u = 1$ and tends to zero as $u \to 1$. Using (4.2.9) with $n = k + 1$, we see that $A_{k+1}(u, \gamma)$ has the same property, whence it follows that it is sufficient to prove the second half of the inequality. But

$$\begin{aligned}
V_1^u A_{k+1}(v,\gamma) &= \int_1^u \left| d_v \int_1^v A_k(v/t,\gamma)\, d_t A_1(t,\gamma)\right| \\
&\leqq \int_1^u \left| d_v \int_1^v |A_k(v/t,\gamma)|\cdot|d_t A_1(t,\gamma)|\right| \\
&\leqq \int_1^u d_v \int_1^v [A_0(v/t,\gamma)]^k\, d_t A_0(t,\gamma) \\
&= \int_1^u [A_0(u/t,\gamma)]^k\, d_t A_0(t,\gamma) \\
&\leqq [A_0(u,\gamma)]^k \int_1^u d_t A_0(t,\gamma) \\
&= [A_0(u,\gamma)]^{k+1}.
\end{aligned}$$

This completes the proof of the inequality.

Let us put

$$(4.2.11) \qquad \begin{cases} B(u,\gamma) = a^{-1} + \sum_{n=1}^{\infty} (-1)^n a^{-n-1} A_n(u,\gamma), & u > 1, \\ B(1,\gamma) = 0. \end{cases}$$

Then by (4.2.10)

$$(4.2.12) \qquad |B(u,\gamma)| \leqq V_1^u B(v,\gamma) \leqq \sum_{n=0}^{\infty} |a|^{-n-1}[A_0(u,\gamma)]^n,$$

the series being absolutely convergent and uniformly bounded for $1 \leqq u \leqq \infty$.

Hence $B(u, \gamma)$ is of bounded variation in $[1, \infty]$ and

$$(4.2.13) \qquad [D(s)]^{-1} = \int_1^\infty u^{-(s-\gamma)}\, d_u B(u,\gamma) = \int_1^\infty u^{-s}\, dB(u),$$

where

$$(4.2.14) \qquad B(u) = \int_1^u v^\gamma\, d_v B(v,\gamma).$$

The second integral in (4.2.13) is clearly absolutely convergent for $\sigma > \max(\alpha, \beta)$. This completes the proof of the theorem.

The assumption that $D(s)$ has a half-plane of absolute convergence is obviously

unnecessarily restrictive. But simple convergence is not enough to insure the existence of even a formal Laplace integral for the reciprocal, much less of a half-plane of convergence.

4.3. *Inversion of the $A(u)$-transform.* In the present paragraph we shall assume that the reciprocal of $D(s)$ admits a representation by means of a convergent Laplace-Stieltjes integral. The functions $A(u)$ and $B(u)$ are then joined by the relation

$$\int_1^u B(u/v)\,dA(v) = \int_1^u A(u/v)\,dB(v) = 1, \qquad u > 1, \tag{4.3.1}$$

for almost all values of u. This is the transcendental analogue of the Möbius algorithm to which it reduces when $A(u)$ is a step function with jumps at the integers.

We can now expect that for a sufficiently restricted class of functions $f(z)$ we shall have

$$\mathfrak{M}\{\mathfrak{M}[f, A], B\} = \mathfrak{M}\{\mathfrak{M}[f, B], A\} = f(z). \tag{4.3.2}$$

We have the following analogue of Theorem 3.2.1.

THEOREM 4.3.1. *A sufficient condition that* (4.3.2) *shall hold is that the double integral*

$$I[f] \equiv \int_1^\infty \int_1^\infty f(uvz)\,dA(u)\,dB(v) \tag{4.3.3}$$

be absolutely convergent.

Proof. Let

$$V_1^u A(t) = A_0(u), \qquad V_1^u B(t) = B_0(u).$$

The condition of the theorem is then that the integral

$$\int_1^\infty \int_1^\infty |\,f(uvz)\,|\,dA_0(u)\,dB_0(v)$$

be convergent. It is then permissible to regard (4.3.3) as a repeated Stieltjes integral and the order in which the integrations are performed is immaterial. Now

$$\mathfrak{M}\{\mathfrak{M}[f, A], B\} = \int_1^\infty \left\{\int_1^\infty f(uvz)\,dA(u)\right\} dB(v),$$

$$\mathfrak{M}\{\mathfrak{M}[f, B], A\} = \int_1^\infty \left\{\int_1^\infty f(uvz)\,dB(v)\right\} dA(u).$$

Hence these two operations exist and give the same result. On the other hand, going back to the definition of the Stieltjes integral as a double sum and "summing by hyperbolas" uv = const. before passing to the limit, we can show that the double integral can be written in the form

$$\int_1^\infty f(wz)\,d_w \int_1^w B(w/u)\,dA(u).$$

Formula (4.3.1) shows that this expression reduces to $f(z)$. This completes the proof of the theorem.

We can now pass to the question of inversion.

THEOREM 4.3.2. *Let $B(u)$ exist as a function of bounded variation and satisfy* (4.3.1). *A sufficient condition that*

$$f(z) = \mathfrak{M}[g, B] \tag{4.3.4}$$

be a solution of

$$g(z) = \mathfrak{M}[f, A] \tag{4.3.5}$$

in the domain S is that the double integral $I[g]$ be absolutely convergent in S. On the other hand, there cannot be more than one solution $f(z)$ of (4.3.5) *which renders $I[f]$ absolutely convergent, and whenever it exists this solution is given by* (4.3.4).

Proof. The assumption that $I[g]$ is absolutely convergent implies that

$$\mathfrak{M}\{\mathfrak{M}[g, B], A\} = g(z)$$

by Theorem 4.3.1. Hence (4.3.4) gives a solution of (4.3.5) and the integral is obviously absolutely convergent.

Conversely, if $f(z)$ is a solution of (4.3.5) such that $I[f]$ is absolutely convergent, then for the same reason

$$f(z) = I[f] = \mathfrak{M}\{\mathfrak{M}[f, A], B\} = \mathfrak{M}[g, B],$$

so that the solution in question is uniquely determined and given by (4.3.4).

Again it is necessary to remark that the mere existence of $\mathfrak{M}[g, B]$ in a domain S is not sufficient to insure that this function be a solution of (4.3.5). Indeed, suppose that the Laplace-Stieltjes integral

$$[D(s)]^{-1} = \int_1^\infty u^{-s}\, dB(u)$$

has a zero $s = \alpha$ in the half-plane of convergence. Then

$$\mathfrak{M}[z^{-\alpha}, B] = z^{-\alpha} \int_1^\infty u^{-\alpha}\, dB(u) \equiv 0$$

and is certainly not a solution of the equation

$$\mathfrak{M}[f(z), A(u)] = z^{-\alpha}.$$

4.4. *Concluding remarks.* In the previous discussion I have perhaps overemphasized the rôle of the associated Laplace-Stieltjes integral

$$D(s) = \int_1^\infty u^{-s}\, dA(u).$$

It should be observed that this function and its reciprocal are mainly tools in the discussion, and the decisive rôle is really played by the reciprocal functions $A(u)$ and $B(u)$ which are supposed to satisfy the algorithm (4.3.1). We know from the sequence case that these functions may very well exist without the

associated Laplace-Stieltjes integrals having any domain of convergence or even any a priori obvious significance. The theorems of §4.3 really presuppose merely the existence of a pair of functions satisfying (4.3.1) and not the existence of the associated Laplace-Stieltjes integrals.

But it must be admitted that when the integrals do not exist, the problem of finding the function $B(u)$ reciprocal to a given function $A(u)$ is in general not a very promising one. Sometimes we may circumvent the difficulties by a preliminary application of a suitable method of summation. Thus it may happen that $s^{-n}[D(s)]^{-1}$ is representable by a Laplace-Stieltjes integral even though $[D(s)]^{-1}$ is not. In this case the corresponding function $B_n(u)$ can be used to find an n-fold integral of the formal solution of (4.3.5) from which the solution itself may be found by solving an integral equation of the Abel type.

There are of course other functional equations which may be treated by the methods of this paper; for instance, the equation

$$\int_{-\infty}^{\infty} F(z - u)\, da(u) = G(z),$$

which is associated with Laplace-Stieltjes integrals having 0 instead of 1 as the lower limit of integration.

Finally, it should be remarked that there are various relations, some obvious, others less so, between the functional equations treated in this paper on one hand, and the theory of Watson transforms and the Karamata-Wiener Tauberian theory on the other. The author hopes to return to these questions at a later opportunity.

YALE UNIVERSITY.

ON BERNOULLI'S NUMBERS AND FERMAT'S LAST THEOREM

By H. S. Vandiver

1. **Introduction.** In the present article a report will be given on the work which has been carried out under a grant made to the writer from the Penrose Fund of the American Philosophical Society.

One of the objects of the work was to extend the known tables of Bernoulli numbers expressed as rational fractions in their lowest terms and to check all previous tables. This was carried out by D. H. Lehmer,[1] who tabulated B_n for $n = 91$ to 110, inclusive. The previous tables had given the values B_n $(n = 1, 2, \cdots, 92)$. The first ninety of these were reproduced in the tables of H. T. Davis.[2] Here

$$B_a = (-1)^{a-1} b_{2a}$$

and

$$(b + 1)^n = b_n \qquad (n \neq 1),$$

where the expression on the left means that $(b + 1)$ is taken to the n-th power by means of the binomial theorem, and b_k is substituted for b^k $(k = 1, 2, \cdots, n)$. All the above mentioned B's were employed by Lehmer in checking the regularity of primes, a prime p being defined as regular if it does not divide the numerators of any of the first $(p - 3)/2$ B's. The details of these latter computations will be treated below.

Another object of the work under the grant was to extend the results of the writer on Fermat's Last Theorem for special exponents.

In other papers[3] it was established that

$$x^l + y^l + z^l = 0 \tag{I}$$

is impossible for x, y, and z non-zero rational integers and l a given integer $2 < l < 307$. In the present article we shall describe the work which established this result for $306 < l < 617$, excepting $l = 587$ which exponent has not yet been tested (cf. note, p. 584). As the methods employed for these large exponents are quite elaborate and complicated, we shall explain many details.

Received May 24, 1937.

[1] This Journal, vol. 2 (1936), pp. 460–464. This article includes references to previous tables. Lehmer computed the value of B_{196} in addition to those mentioned but not as part of the present project. Cf. Annals of Math., vol. 36 (1935), p. 648.

[2] *Tables of Higher Mathematical Functions*, vol. 2, Bloomington, Indiana, 1935, pp. 230–233.

[3] Proc. Natl. Acad. Sci., vol. 17 (1931), pp. 661–673 (referred to later as N. A.) with references there given.

We have persisted in the examination of special exponents in (I) in the hope that, if one of the criteria that we have employed throughout (N. A., p. 670, Theorem IV) for irregular primes l breaks down for a particular l, we shall find such an l in the range of our computations. So far no exponent to which we have applied the criterion has belonged to this class.

2. **Congruence properties of the Bernoulli numbers.** In a previous paper a number of congruences of this character were derived,[4] and we shall here generalize and simplify some of these results and derive others to be employed later. If p is prime, consider the identity

$$(x^{pk} - 1)\frac{x^{pm} - 1}{x - 1} = (x^{pm} - 1)\frac{x^{pk} - 1}{x - 1}, \tag{1}$$

and write

$$f_a^{(s)}(x) = x + 2^{a-1}x^2 + \cdots + (ps - 1)^{a-1}x^{ps-1}$$

for $a > 1$ and

$$f^{(s)}(x) = 1 + x + x^2 + \cdots + x^{ps-1}.$$

These functions will be called *Mirimanoff polynomials*. If e is the Napierian base, we see that

$$\left[\frac{df_a^{(s)}(e^v x)}{dv}\right]_{v=0} = f_{a+1}^{(s)}(x). \tag{2}$$

In lieu of this we may employ the formal operation

$$\frac{x\,df_a^{(s)}(x)}{dx},$$

but we shall use the exponential function as most of the papers which have been written along these lines employ it.

Setting $x = e^v x$ in (1), differentiating a times, and setting $v = 0$, we have, employing Leibnitz' theorem,

$$\begin{aligned}(x^{pk} - 1)f_{a+1}^{(m)}(x) &+ apkx^{pk}f_a^{(m)}(x)\\ &\equiv (x^{pm} - 1)f_{a+1}^{(k)}(x) + apmx^{pm}f_a^{(k)}(x) \pmod{p^2}.\end{aligned} \tag{3}$$

Let ρ be an m-th root of unity $\neq 1$; then the last relation gives, if $m \not\equiv 0 \pmod{p}$,

$$\sum_{k=1}^{m-1}(\rho^{pk} - 1)f_{a+1}^{(m)}(\rho) + ap\sum_{k=1}^{m-1}k\rho^{pk}f_a^{(m)}(\rho) \equiv apm\sum_{k=1}^{m-1}f^{(k)}(\rho) \pmod{p^2}.$$

Now

$$f_a^{(m)}(\rho) \equiv (1 + \rho^p + \cdots + \rho^{(m-1)p})f_a^{(1)}(\rho) \equiv 0 \pmod{p},$$

[4] Proc. Natl. Acad. Sci., vol. 16 (1930), pp. 139–150.

whence

$$(4) \qquad -\sum{}' m f_{a+1}^{(m)}(\rho) \equiv apm \sum_{k=1}^{m-1} \sum{}' f_a^{(k)}(\rho) \qquad (\text{mod } p^2),$$

where $\sum'$ indicates summation over all the distinct values of ρ.

Consider the summation

$$\sum{}' f_{a+1}^{(m)}(\rho).$$

If we carry the summation over each term, then the term of the form $\sum' t^a \rho^t = -t^a$ for t not divisible by m, whereas if t is divisible by m the term is $(m-1)\, t^a$. Noting that if $a \not\equiv 1 \pmod{p-1}$,

$$\sum_{i=1}^{pm-1} i^a \equiv \sum_{h=0}^{m-1} \sum_{j=0}^{p-1} (j+ph)^a \equiv mS_a + \sum_h aph\, S_{a-1} \ (\text{mod } p^2); \quad S_a = \sum_{i=1}^{p-1} i^a,$$

we obtain, using $S_{a-1} \equiv 0 \pmod p$, the relation

$$(5) \qquad \sum{}' f_{a+1}^{(m)}(\rho) \equiv (m^{a+1} - m) S_a \qquad (\text{mod } p^2).$$

Hence (4), (5), and the known relation $S_a \equiv pb_a \pmod{p^2}$ give for a even, $p > 3$,

$$(6) \qquad -\frac{m^{a+1} - m}{a} b_a \equiv \sum_{k=1}^{m-1} \sum{}' f_a^{(k)}(\rho) \qquad (\text{mod } p),$$

whence

$$(7) \qquad \frac{1 - m^a}{am^{a-1}} b_a \equiv \sum_{v=1}^{m-1} \sum_{j=1}^{[vp/m]} j^{a-1} \qquad (\text{mod } p),$$

where $[x]$ is the greatest integer in x.

We have

$$\left[\frac{(n-k)p}{n}\right] = \left[p - \frac{kp}{n}\right] = p - \left[\frac{kp}{n}\right] - 1,$$

whence

$$(8) \qquad \left(\left[\frac{(n-k)p}{n}\right] - i + 1\right)^{2a-1} \equiv -\left(\left[\frac{kp}{n}\right] + i\right)^{2a-1} \qquad (\text{mod } p).$$

From this, we may show that, if $k < n$,

$$(9) \qquad \begin{aligned} C_{k+1}^{(n)} + C_{n-k}^{(n)} &\equiv 0 \qquad (\text{mod } p); \\ C_l^{(m)} &= \sum_{i=[(l-1)p/m]+1}^{[lp/m]} i^{2n-1}. \end{aligned}$$

In $C_{k+1}^{(n)}$ there are

$$\left[\frac{(k+1)p}{n}\right] - \left[\frac{kp}{n}\right] = s$$

terms, and in $C_{n-k}^{(n)}$ there are

$$\left[\frac{(n-k)p}{n}\right]-\left[\frac{(n-k-1)p}{n}\right]$$

terms, and these two numbers are equal since

$$\left[\frac{(n-k)p}{n}\right]=\left[p-\frac{kp}{n}\right]=p-\left[\frac{kp}{n}\right]-1$$

with a similar relation in which $k+1$ replaces k. Now (8) gives

$$\sum_{i=1}^{s}\left(\left[\frac{(n-k)p}{n}\right]-i+1\right)^{2a-1}\equiv-\sum_{i=1}^{s}\left(\left[\frac{kp}{n}\right]+i\right)^{2a-1}\pmod{p}.$$

This is the relation (9). Now consider

$$(n-l)C_l^{(n)}+(l-1)C_{n+1-l}^{(n)}.$$

Using (9) we see that this is congruent modulo p to

$$(n-2l+1)C_l^{(n)}.$$

Also for n odd we have

$$C_{(n+1)/2}^{(n)}\equiv 0\pmod{p}$$

from (9). Hence from (9) we have, using (7),

$$(10)\qquad \begin{aligned} b_{2a}\frac{1-n^{2a}}{2an^{2a-1}}&\equiv\sum_{l=1}^{[n/2]}(n-2l+1)C_l^{(n)}\pmod{p}\\ \sum_{l=1}^{m-1}(m-l)C_l^{(m)}&=\sum_{l=1}^{m-1}\sum_{\mu=1}^{l}C_\mu^{(m)}.\end{aligned}$$

For an odd $n=n_1$ this reduces to

$$(10a)\qquad b_{2a}\frac{1-n_1^{2a}}{4an_1^{2a-1}}\equiv\sum_{l=1}^{[n_1/2]}\frac{(n_1+1-2l)}{2}C_l^{(n_1)}\pmod{p}.$$

To accord with the notation employed in other papers we write

$$C_l^{(6)}=C_l;\qquad C_l^{(4)}=A_l$$

defined for $p>6$.

Now for $2a<(p-1)$ we have

$$\begin{aligned} b_{2a}\frac{1-n^{2a}}{2an^{2a-1}}&\equiv b_{2a}\frac{n^{p-1}-n^{2a}}{2an^{2a-1}}\\ &\equiv b_{2a}\frac{n^{p-2a}-n}{2a}\pmod{p}.\end{aligned}$$

Hence we obtain from (10) and (10a)

$$(11) \qquad \frac{(6^{p-2a}-6)b_{2a}}{2a} \equiv 5C_1 + 3C_2 + C_3 \pmod{p};$$

$$(12) \qquad \frac{(4^{p-2a}-4)b_{2a}}{2a} \equiv 3A_1 + A_2 \pmod{p};$$

$$(13) \qquad \frac{(3^{p-2a}-3)b_{2a}}{4a} \equiv C_1 + C_2 \pmod{p},$$

observing that

$$C_1 + C_2 = \sum_{s=1}^{[p/3]} s^{2a-1}.$$

We note also that

$$A_1 + A_2 = C_1 + C_2 + C_3$$

and (11) gives

$$\frac{(6^{p-2a}-6)b_{2a}}{2a} \equiv 2(C_1+C_2) + 2C_1 + A_1 + A_2$$

$$\equiv \frac{3^{p-2a}-3}{2a} b_{2a} + \frac{4^{p-2a}-4}{2a} b_{2a} + 2(C_1 - A_1) \pmod{p}.$$

This gives

$$A_1 - C_1 \equiv (-1)^a B_a \frac{1 - 3^{p-2a} - 4^{p-2a} + 6^{p-2a}}{4a}$$

$$\equiv (-1)^a B_a \frac{(2^{p-2a}-1)(3^{p-2a}-2^{p-2a}-1)}{4a} \pmod{p},$$

and this may be put in the form

$$(14) \qquad \sum_{s=[p/6]+1}^{[p/4]} s^{2a-1} \equiv (-1)^a B_a \frac{(2^{p-2a}-1)(3^{p-2a}-2^{p-2a}-1)}{4a} \pmod{p}$$

for $p > 7$, $2a < p - 1$.

From (10) we have

$$(15) \qquad \frac{(2^{p-2a}-2)b_{2a}}{2a} \equiv A_1 + A_2 \pmod{p}$$

which with (13) gives

$$\frac{3^{p-2a}-3}{2a} b_{2a} \equiv 2(A_1 + A_2) - 2C_3 \pmod{p}$$

or

$$(16) \qquad \sum_{[l/3]+1}^{[l/2]} s^{2a-1} \equiv \frac{2^{l-2a+1} - 3^{l-2a} - 1}{4a} (-1)^{a-1} B_a \pmod{p}.$$

Now set

$$D_h = C_h^{(5)},$$

then (10a) gives

$$b_{2a} \frac{1 - 5^{2a}}{4a \cdot 5^{2a-1}} \equiv 2D_1 + D_2 \pmod{p}. \tag{17}$$

Now (11) and (15) give

$$\frac{(6^{p-2a} - 6)b_{2a}}{2a} \equiv 4C_1 + 2C_2 + \frac{(2^{p-2a} - 2)b_{2a}}{2a} \pmod{p}$$

or

$$\frac{b_{2a}(6^{p-2a} - 2^{p-2a} - 4)}{4a} \equiv 2C_1 + C_2 \pmod{p}.$$

Subtraction of this from (17) gives, modulo p,

$$2(D_1 - C_1) + D_2 - C_2 \equiv b_{2a} \frac{5^{p-2a} - 6^{p-2a} + 2^{p-2a} - 1}{4a}. \tag{18}$$

3. **Examination of primes as to regularity.** The first step in our examination of a particular l in (I) was to determine if it was regular, so that we could apply the known result of Kummer to the effect that (I) is impossible if l is a regular prime. For this purpose formulas (14) and (16) were mainly employed heretofore. For the larger primes, however, (18) was found to be more valuable and was mainly employed in the present work, where $l = p$. The number of values of s used in (14) is approximately $[l - l/12]$. The number required in (18) is larger, but there exists a relation between the values of s in (18) which does not hold in (14). The range of values for s in the expression $2(D_1 - C_1)$ is from $[l/6] + 1$ to $[l/5]$, where each value is used twice. The expression $D_2 - C_2$ contains the values from $[l/5] + 1$ to $[2l/5]$ less the values from $[l/6] + 1$ to $[l/3]$. Thus the negative values in the second expression cancel the double values of the first, leaving the values of s in (18) in two ranges, $[l/6] + 1$ to $[l/5]$ and $[l/3] + 1$ to $[2l/5]$.

Let n be a number in the range $[l/6] + 1$ to $[l/5]$. Then $n = [l/6] + i$, where i is a positive integer. Let $j = [l/6]$. Then $l/6 = j + k/6$, where $0 < k < 6$, and $l/3 = 2j + k/3$. Therefore, $2[l/6] \geq [l/3] - 1$, and $2n \geq [l/3] + 1$. Also $n \leq [l/5]$, and as above $2[l/5] \leq [2l/5]$. Therefore, $2n \leq [2l/5]$. Thus we have the relation:

(19) *If n is a number in the range $[l/6] + 1$ to $[l/5]$, then $2n$ is in the range*

$$[l/3] + 1 \text{ to } [2l/5].$$

It also follows from the above argument that if n is an integer such that $2n$ is in the range $[l/3] + 1$ to $[2l/5]$, then n is in the range $[l/6] + 1$ to $[l/5]$. Thus for a particular value of a in (18) it is sufficient to compute the powers of the

odd values of s in the range $[l/3] + 1$ to $[2l/5]$. The sums of the powers of the even values may be found in one operation by computing

$$2^{2a-1} \sum_{s=[l/6]+1}^{[l/5]} s^{2a-1}. \tag{20}$$

M. M. Abernathy has obtained a result concerning primes of the form $4n + 1$. If index $s = 2k$, then

$$\text{ind } s^{(l-1)/2} = 2n \cdot 2k = 4nk \equiv 0 \pmod{\overline{l-1}}.$$

Therefore $s^{(l-1)/2} \equiv 1 \pmod{l}$ and $s^{(l-1)/2+i} \equiv s^i \pmod{l}$, where i and k are integers and l is a given prime of the form $4n + 1$. If index $s = 2k + 1$, then

$$\text{ind } s^{(l-1)/2} = 2n(2k + 1) = 4nk + 2n \equiv 2n \pmod{\overline{l-1}}.$$

Therefore $s^{(l-1)/2} \equiv -1 \pmod{l}$ and $s^{(l-1)/2+i} \equiv -s^i \pmod{l}$. Thus in testing for regularity of primes of the form $4n + 1$, powers of s greater than $(l - 1)/2$ do not have to be computed. Division of values of s into two groups, those of even and those of odd index, permits computation of

$$\sum_{s=i}^{j} s^{2a-1}, \qquad a > \frac{l-1}{4}$$

by taking

$$\sum_{s_1} s_1^{2a-1-(l-1)/2} - \sum_{s_2} s_2^{2a-1-(l-1)/2}, \tag{21}$$

where s_1 ranges over values with even index, s_2 over those of odd index.

In computing the right-hand member of (18) it sometimes happened that this was $\equiv 0 \pmod{l}$ because the second factor on the right had its numerator divisible by $l = p$. In these cases (14), (15), and (16) were employed until the numerator of the factor not involving b_{2a} was found to be $\not\equiv 0 \pmod{l}$. When a B_a was found with its numerator divisible by l, then this was checked by employing the tables of Bernoulli numbers of Lehmer or of previous writers, provided the Bernoulli number fell within the range of their computations. Further, to aid in this part of the work, Lehmer suggested and carried out the division of each B_a ($a \leq 110$) by each of the primes l, where $547 \leq l < 601$, and for all primes l_1 of the form $4n + 3$, where $601 < l_1 < 619$. Jacobi's table of indices[5] was employed to carry out the other part of the computations. We shall now give an example of the method for $l = p = 541$. For this case, the terms in the right-hand member of (18) are

$$2(D_1 - C_1) = 2 \sum_{s=91}^{108} s^{2a-1},$$

$$(D_2 - C_2) = \sum_{s=181}^{216} s^{2a-1}.$$

[5] *Canon Arithmeticus*, Berlin, 1839.

For each s appearing in the above ranges, s^{2a-1} was computed employing Jacobi's tables for $a = 1, 2, \cdots, (l - 3)/2$; then the computations of $D_1 - C_1$ and $D_2 - C_2$ were each separated into two parts since information about some of the terms s^{2a-1} in the second sum can be obtained immediately from some of those of the first, as we indicated at the beginning of §3. For the primes mentioned above tested by Lehmer we started with $a = 111$ in lieu of $a = 1$.

In the computation of s^{2a-1} $(a = 1, 2, \cdots, (l - 3)/2)$, periodicities may appear, which, since we are using indices, may be determined in advance. If $l - 1 = km$ and index $s = kn$, then s will begin repeating at s^{m+1}; for $s^{m+1} \equiv s \pmod{l}$ follows from the fact that since $r^{kn} \equiv s$,

$$kn(m + 1) = kmn + kn \equiv kn \qquad (\mathrm{mod}\ \overline{l - 1}).$$

At any stage of the work we have another check by the possible use of the formula

$$\sum_{a=1}^{k} s^{2a-1} = \frac{s^{2k+1} - s}{s^2 - 1}$$

and in particular the convenient relation

$$\sum_{a=1}^{(l-1)/2} s^{2a-1} \equiv 0 \qquad (\mathrm{mod}\ l).$$

The regular[6] primes l $(306 < l < 619)$ are as follows: 313, 317, 331, 337, 349, 359, 367, 373, 383, 389, 397, 419, 431, 439, 443, 449, 457, 479, 487, 499, 503, 509, 521, 563, 569, 571, 599, 601, 613, and the irregular primes are 307, 311, 347, 353, 379, 401, 409, 421, 433, 461, 463, 467, 491, 523, 541, 547, 557, 577, 587, 593, 607, 617.

As to the time required for the work on a particular prime l, close to the value 600, in a test for regularity, a person experienced in this work requires on the average about forty hours to complete the test provided the prime is of the form $4n + 3$. A prime of the form $4n + 1$ requires about two thirds of this time.

Concerning the numbers in the set

$$B_1, B_2, \cdots, B_{(l-3)/2} \tag{22}$$

which are divisible by l for a particular l we found

$$B_n \equiv 0 \qquad (\mathrm{mod}\ l)$$

in the following cases: $l = 307, n = 44$; $l = 311, n = 146$; $l = 347, n = 140$; $l = 353, n = 93, n = 150$; $l = 379, n = 50, n = 87$; $l = 401, n = 191$; $l = 409, n = 63$; $l = 421, n = 120$; $l = 433, n = 183$; $l = 461, n = 98$; $l = 463, n = 65$; $l = 467, n = 47, n = 97$; $l = 491, n = 146, n = 169$; $l = 523, n = 123, n = 200$; $l = 541, n = 43$; $l = 547, n = 135, n = 243$; $l = 557, n = 111$; $l = 577, n = 26$; $l = 587, n = 45, n = 46$; $l = 593, n = 11$; $l = 607, n = 296$; $l = 617, n = 10, n = 87, n = 169$. From the above it will be noted that just two of the B's in the set (22) are divisible by l in each of the cases $l = 353, 379,$

[6] The irregular primes < 211 are listed by the writer in the Transactions of the American Mathematical Society, vol. 31, pp. 613, 615–616, and for $211 \leq l < 307$ in N. A., p. 667.

467, 491, 523, 547, 587. For 617, three of the B's in (22) are divisible by 617 and this is the first l of this sort encountered in our work since the beginning.

4. **Treatment of the exponents in (I) which are irregular primes.** Here we employ the following:[7]

THEOREM 1. *Under the assumptions: none of the units* E_a $(a = a_1, a_2, \cdots, a_s)$ *is congruent to the l-th power of an integer in* $k(\zeta)$ *modulo* $\mathfrak{P}$, *where* $\mathfrak{P}$ *is a prime ideal divisor of* p; p *is a prime* $< (l^2 - l)$ *of the form* $1 + lk$; *and* $a_1, a_2, \cdots, a_s$ *are the subscripts of the B's in the set*

$$B_1, B_2, \cdots, B_{(l-3)/2}$$

whose numerators are divisible by l; *the relation* (I) *is impossible in non-zero rational integers* x, y, *and* z.

In the above statement,

$$E_n = \prod_{i=0}^{(l-3)/2} \epsilon(\zeta^{r^i})^{r^{-2in}};$$

$$\epsilon = \left(\frac{(1 - \zeta^r)(1 - \zeta^{-r})}{(1 - \zeta)(1 - \zeta^{-1})}\right)^{\frac{1}{2}};$$

r being a primitive root of l; $\zeta = e^{2i\pi/l}$. As indicated in N. A., p. 667, to test this for a particular irregular l and a $p < (l^2 - l)$ of the form $1 + kl$, we compute the value of $E_n(d)$ modulo p, where $E_n(d)$ is written in the form

$$d^R \prod_{i=0}^{(l-3)/2} \left(\frac{d^{r^{i+1}} - 1}{d^{r^i} - 1}\right)^{r^{(h-1)i}},$$

$$R = \frac{1 - r}{2}(1 + r^h + r^{2h} + \cdots + r^{\frac{1}{2}(l-3)h}),$$

$$h = l - 2n.$$

d is a rational integer such that

$$d^l \equiv 1 \pmod{p}.$$

In N. A., p. 668 a systematic method of carrying out such a computation was explained for the case $l = 271$, $p = 1627$. M. E. Tittle has further abbreviated this scheme in general. Since

$$d^l \equiv 1 \pmod{p},$$

all values of $(d^{r^i} - 1)$ are obtained immediately from a table of the first powers of d, modulo p. This renders unnecessary the third row employed in N. A., p. 668 and eliminates the computation involved in obtaining the fourth row. A companion table to the powers of d, with powers of ρ, a primitive root of p, substituted and with blank spaces for the omitted numbers permits calculation of ind $(d^{r^i} - 1)$. If $(d^{r^i} - 1)$ is not one of the numbers appearing in the

[7] N. A., p. 670, Theorem IV; Bulletin of the American Mathematical Society, vol. 40 (1934), p. 124, paragraph immediately following statement of Theorem 2.

companion table, multiply $(d^{r^i} - 1)$ by ρ, and examine the table for the appearance of this number. Repetition of this operation will yield a result in at most $k - 1$ steps, for the numbers $n, \rho n, \rho^2 n, \cdots, \rho^{k-1}n$, have as indices the numbers $m, m + 1, m + 2, \cdots, m + k - 1$, where index $n = m$, and one of the numbers in this last set is of the form ki, i an integer, and all indices of the form ki are presented. To obtain the actual value of $(d^{r^i} - 1)$, subtract from the observed value in the table the power of ρ required in the multiplication $\rho^j(d^{r^i} - 1)$. Detailed checks on this type of work are described in N. A., p. 669.

The following table gives the specific results of the computation for each irregular prime. In the table, l is the prime exponent appearing in (I), n is a Bernoulli number in the set (22) whose numerator is divisible by l, r is a primitive root of l, d is the integer selected such that

$$d^l \equiv 1 \pmod{p},$$

p is a prime of the form $1 + kl$ referred to in the statement of Theorem 1, ρ is a primitive root of p, ind is the index, modulo l, found for $E_n(d)$, modulo p. When two values of n are listed for a particular l, then the corresponding indices are listed in the same order in the last column. Thus for $l = 353$, $B_{93} \equiv 0 \pmod{353}$, ind $E_{93}(2804) \equiv 57 \pmod{353}$; $B_{150} \equiv 0 \pmod{353}$, ind $E_{150}(2804) \equiv 13 \pmod{353}$. The exponents $l = 587$ and $l = 617$ have not yet been tested by the criteria of Theorem 1.

l	n	r	d	p	ρ	ind
307	44	5	168	1229	10	213
311	146	308	135	1867	1857	27
347	140	337	64	2083	2	333
353	93, 150	3	2804	4943	10	57, 13
379	50, 87	2	3954	4549	6	101, 200
401	191	211	3^8	3209	3	365
409	63	235	2^4	1637	2	115
421	120	238	6^{10}	4211	6	338
433	183	10	2^4	1733	2	332
461	98	10	10^6	2767	10	190
463	65	174	2^{12}	5557	2	215
467	47, 97	−10	2^6	2803	2	347, 87
491	146, 169	10	5^2	983	5	148, 382
523	123, 200	−10	2^{10}	5231	2	85, 228
541	43	10	3^{18}	9739	3	458
547	135, 243	17	3^{10}	5471	−3	477, 179
557	111	41	10^6	3343	10	222
577	26	10	15	2309	2	556
587	45, 46					
593	11	10	2^2	1187	2	523
607	296	575	2^6	3643	2	491
617	10, 87, 169					

It is a little curious that the smallest prime p satisfying the conditions in Theorem 1 gave for each l

$$\text{ind } E_n(d) \not\equiv 0 \pmod{l}.$$

We may now state

THEOREM 2.

$$x^n + y^n = z^n$$

is impossible in non-zero rational integers x, y, and z if n is a given integer, $2 < n < 617$, excepting possibly 587.

The above results have also been established for various prime exponents $l < 700$ not included in the above. The computations are proceeding under other auspices.

5. **Application of the data obtained to other parts of the theory of cyclotomic fields.** The numerical data obtained gives special information concerning many other questions in cyclotomic field theory. First, various results are known[8] concerning regular cyclotomic fields which have not been extended to other fields, and which have been applied to obtain results in Diophantine Analysis. By means of our work described here, we have isolated many regular cyclotomic fields in which all the results just referred to apply.

Let us now consider *primary units* in a cyclotomic field. A unit η is primary in the field defined by a primitive l-th root of unity $k(\zeta)$ if it is not the l-th power of an integer in $k(\zeta)$ and if

$$\eta \equiv \alpha^l \pmod{\lambda^l},$$

where α is an integer in $k(\zeta)$ and $\lambda = (1 - \zeta)$. If l is regular[9] there is no primary unit in $k(\zeta)$. If l is irregular, we may show that $E_n(\zeta)$ is primary provided $B_n \equiv 0 \pmod{l}$. For, we may set

$$E_n(\zeta) = E_n(1) + \theta_1(\zeta)(1 - \zeta)^a.$$

Then for an indeterminate w

$$w^{l^2} E_n(w)^{l-1} = E_n(1)^{l-1} + \theta(w)(1 - w)^a + V(w)\left(\frac{w^l - 1}{w - 1}\right) \tag{23}$$

since $V(w)$ is some polynomial in w with rational integral coefficients. In this relation set $w = e^v$, take the logarithms of each member, differentiate k times, and set $v = 0$. Then[10] since for $i \neq n$

$$\left[\frac{d^{2i} \log E_n(e^v)}{dv^{2i}}\right]_{v=0} = \frac{r^{(l-1)(i-n)} - 1}{r^{2i-2n} - 1} (-1)^{i+1} \frac{B_i}{2i} (r^{2i} - 1), \tag{24}$$

[8] Hilbert, *Werke*, I, pp. 278–312. Maillet, Annali di Mat., (3), vol. 12 (1906), pp. 145–178; Acta Math., vol. 24 (1901), pp. 247–256; Vandiver, Proc. Natl. Acad. Sci., vol. 17 (1931), pp. 662–663; Monats. Math. u. Phys., vol. 43 (1936), pp. 317–320.

[9] Hilbert, loc. cit., p. 287.

[10] Vandiver, Transactions of the American Mathematical Society, vol. 31 (1929), pp. 619–620.

and for $i = n$

$$\left[\frac{d^{2i}\log E_n(e^v)}{dv^{2n}}\right]_{v=0} = \frac{(-1)^{n+1}B_n(l-1)(r^{2n}-1)}{4n},$$

using $B_n = 0$ with

$$\left[\frac{d^{2i+1}\log E_n(e^v)}{dv^{2i+1}}\right]_{v=0} = 0,$$

we obtain

$$\left[\frac{d^k\theta(e^v)(1-e^v)^a}{dv^k}\right] \equiv 0 \pmod{l} \qquad (k = 1, 2, \cdots, l-2).$$

Hence

$$\theta_1(\zeta)(1-\zeta)^a = \chi(\zeta)(1-\zeta)^{l-1}.$$

To prove that $\chi(\zeta)$ is divisible by $1 - \zeta$ and hence that $E_n(\zeta)$ is primary we may write (23) in the form, d being some integer,

$$(25) \quad w^{l^2}E_n(w)^{l-1} = E_n(1)^{l-1} + \chi(w)(1-w)^{l-1} + V_1(w)(w^l - 1) + d\,\frac{w^l-1}{w-1}.$$

In this relation set $w = 1$; this gives $d = 0$. In (25) set $w = e^v$, take logarithms of each member, differentiate $(l-1)$ times, and set $v = 0$, and we have, employing (24), since $\left[\frac{d^s(1-e^v)^{l-1}}{dv^s}\right]_{v=0} \equiv 0 \pmod{l}$ $(s < l-1)$,

$$\left[\frac{d^{l-1}\chi(e^v)(1-e^v)^{l-1}}{dv^{l-1}}\right]_{v=0} \equiv 0 \pmod{l},$$

and

$$\left[\chi(e^v)\frac{d^{l-1}(1-e^v)^{l-1}}{dv^{l-1}}\right]_{v=0} \equiv 0 \pmod{l},$$

whence

$$[\chi(e^v)]_{v=0} \equiv 0 \pmod{l},$$

or

$$\chi(e^v) \equiv (1-e^v)w(e^v) \pmod{l}.$$

This gives the result.

We may now show that $E_n(\zeta)$ is primary when $B_n \equiv 0 \pmod{l}$ if we note also that E_n is never the l-th power of a unit in $k(\zeta)$ in any of the cases we have tested since in each such case we found a prime ideal $\mathfrak{P}$ such that $E_n(\zeta)$ is not congruent to the l-th power of an integer in $k(\zeta)$ modulo $\mathfrak{P}$, where $\mathfrak{P}$ is a prime ideal divisor of p. Hence for any l where we have found $E_n(\zeta)$ primary we have explicitly determined an absolute (Hilbert) class field[11] of $k(\zeta)$.

[11] Hilbert, loc. cit., pp. 149–156.

We now consider the connection of our results with the second factor of the class number of $k(\zeta)$. This factor may be written[12]

$$\frac{n_1 n_2 \cdots n_{l_1}}{\Delta},$$

where Δ is the determinant

$$\begin{vmatrix} b_{11} & b_{21} & \cdots & b_{l_1 1} \\ b_{12} & b_{22} & \cdots & b_{l_1 2} \\ \cdots & \cdots & \cdots & \cdots \\ b_{1 l_1} & b_{2 l_1} & \cdots & b_{l_1 l_1} \end{vmatrix},$$

the b's and n's being rational integers not all zero such that

$$\gamma_s^{n_s} = \epsilon_1^{b_{1s}} \epsilon_2^{b_{2s}} \cdots \epsilon_{l_1}^{b_{l_1 s}} \qquad (s = 1, 2, \cdots, l_1),$$

where $\gamma_1, \gamma_2, \cdots, \gamma_{l_1}$ is a set of fundamental units of $k(\zeta)$; ϵ_i is the unit obtained from our ϵ, previously defined, by the substitution $(\zeta/\zeta^{r^{i-1}})$; $l_1 = (l-3)/2$. In a former paper,[13] the writer showed that for this second factor (say h_2) to be divisible by l it is necessary and sufficient that at least one of the units $E_i(\zeta)$ $(i = 1, 2, \cdots, l_1)$ be the l-th power of a unit in $k(\zeta)$. In view of what we observed above, concerning the E's not being congruent to the l-th power of an integer in $k(\zeta)$ modulo $\mathfrak{P}$, we may state the

THEOREM 3. *If l is an odd prime and $\zeta = e^{2i\pi/l}$, then the second factor of the class number of the field $k(\zeta)$ is prime to l for each $l < 617$ excepting possibly* 587.

If h is the class number of $k(\zeta)$, where $k(\zeta)$ is defined by an irregular prime, then we may write

$$h = l^t j,$$

where j is prime to l. If we raise each ideal in $k(\zeta)$ to the j-th power, then the ideal classes defined by these ideals will form a group called the irregular class group of $k(\zeta)$. The prime ideal $\mathfrak{P}$, the factor of (p), which we have used in connection with each E_n which was primary, is such that $\mathfrak{P}^j$ belongs to this irregular class group. This follows since $E_n(\zeta)$ is not congruent to the l-th power of an integer in the field $k(\zeta)$, and, since $E_n(\zeta)$ is primary, this cannot hold if $\mathfrak{P}$ does not belong to the irregular class group. That is, $\mathfrak{P}^k$ is not a principal ideal unless k is divisible by l. For example, turning to the table we see that for $l = 307$ the prime ideal $\mathfrak{P}$, a prime ideal divisor of 1229, belongs to the irregular class group, etc. Specifically, such ideals can be written in the form

$$(\zeta - \rho^{(p-1)/l}, p)$$

and since they are not principal, they cannot be further reduced.

[12] Report on the Theory of Algebraic Numbers, Bull. Nat. Res. Coun., No. 2, February, 1928, pp. 34 and 38, with references to Kummer there given; Fueter, *Synthetische Zahlentheorie*, Berlin and Leipzig, 1925, 2d ed., p. 223.

[13] Proc. Natl. Acad. Sci., vol. 16 (1930), pp. 743–749.

In connection with some previous work on Fermat's Last Theorem we employed for particular exponents the following (N. A., p. 670)

THEOREM 4. *Under the following assumptions*

(1) *the second factor of the class number of the field $k(\zeta)$ is prime to l;*

(2) *none of the Bernoulli numbers B_{nl} $(n = 1, 2, \cdots)$, $(l - 3)/2$, is divisible by l^3,*

the equation (I) *is impossible in rational non-zero integers x, y, and z.*

The above theorem was applied to prove Fermat's Last Theorem for all primes $l < 307$. We are enabled to remove the restriction to case II mentioned in the theorem in N. A. because of the results in another paper of the writer's.[14] The question involved in the second assumption in the theorem is intimately connected with other questions concerning the divisibility of certain other Bernoulli numbers by l^2. We have the known relation

$$B'_{n+2\mu} - 2B'_{n+\mu} + B'_n \equiv 0 \pmod{l^2}; \qquad \mu = (l-1)/2,$$

where

$$B'_i = \frac{(-1)^i B_i}{i}.$$

From this we obtain easily by induction[15]

$$B'_{n+y\mu} \equiv yB'_{n+\mu} - (y-1)B'_n \pmod{l^2}. \tag{26}$$

If we now select l and n so that $B'_n \equiv 0 \pmod{l}$ with $B'_{n+\mu} \not\equiv B'_n \pmod{l^2}$, it follows that there exists an integer y which yields $B'_{n+y\mu} \equiv 0 \pmod{l^2}$. Pollaczek took $l = 37$, $n = 16$, which gave $y = 7$; hence $B'_{142} \equiv 0 \pmod{37^2}$. He also found two other examples of this type. Hence the numerator of a Bernoulli number may be divisible by the square of a proper divisor, where a proper divisor of the numerator of a Bernoulli number B_i is one which is prime to i.

Pollaczek showed that a necessary condition that $k(\zeta)$ contain an ideal belonging to the exponent l^2 was that one of the $(l - 3)/2$ Bernoulli numbers $B_{(tl+1)/2}$ $(t = 1, 3, \cdots, l - 4)$ be divisible by l^2. In view of the above, we consider the number of Bernoulli numbers in the set

$$B_n, B_{n+\mu}, \cdots, B_{n+(l-1)\mu}$$

which are divisible by l^2. The relation (26) shows that if two of these numbers are divisible by l^2, then all of them are. Here we are assuming $B_n \equiv 0 \pmod{l}$. The work of Pollaczek mentioned above for the case $l = 37$ shows that

$$B'_{n+\mu} \not\equiv B'_n \pmod{l^2}.$$

Hence it follows that, since

$$B'_{142} \equiv 0 \pmod{37^2},$$

$$\frac{B_{nl}}{l} \not\equiv 0 \pmod{l^2}$$

[14] Bulletin of the American Mathematical Society, vol. 40 (1934), p. 118, Theorem 1.

[15] Pollaczek, Math. Zeitschrift, vol. 21 (1924), p. 36.

or

$$B_{nl} \not\equiv 0 \pmod{l^3}$$

for $n = 16, l = 37$ likewise

$$B_{nl-\mu} \not\equiv 0 \pmod{l^2}.$$

Here we note that

$$nl - \mu = \frac{(2n-1)l + 1}{2},$$

and if

$$B_{nl-\mu} \equiv 0 \pmod{l^2},$$

we may have an ideal $k(\zeta)$ belonging to the exponent l^2. Hence Pollaczek's computations verify independently for $l = 37$ that the class number of $k(\zeta)$ is divisible by l but not by l^2 and also that

$$N_{16\cdot 37} \not\equiv 0 \pmod{37^3}.$$

His work (loc. cit.) in connection with the primes 59 and 67 furnishes similar checks.

The data obtained concerning the examination of primes as to regularity furnishes much other information concerning the divisibility properties of the Bernoulli numbers. For example, we may find the least positive residue of the numerator of any Bernoulli number modulo l, where l is any prime < 619. For, in each case we have computed the left-hand members of one of the formulas (14), (15), (16), or (18), finding the least positive residue modulo l for the p used in these formulas. Thus in (14) we may conveniently reduce the factor on the right not involving B_a employing Jacobi's table of indices and also using the explicit known form of the denominator of B_a. For $a > (l-3)/2$ we employ the known formula

$$(-1)^{\mu} \frac{B_{a-\mu}}{a-\mu} \equiv \frac{B_a}{a} \pmod{l}$$

except when a is a multiple of μ, $\mu = (l-1)/2$. In the latter case we have

$$pB_{k\mu} \equiv -1 \pmod{l}$$

which enables us to reduce the numerator of $B_{k\mu}$, modulo l.

We may note that arithmetical machines were not employed in the computations described in this paper except in connection with addition. In my opinion, there are probably many other types of number theoretic computation in which, in the future, it will be very convenient to employ, if they exist, tables of indices for primes beyond 1000. In connection with the second type of computation involving the treatment of irregular exponents in (1) we constructed special tables of indices as previously described. That is, for the value of p which we used, we set up partial companion tables from which we can quickly obtain the smallest residue of any integer raised to any power modulo p and having

given an integer we can find what power of ρ is congruent to that integer modulo p where ρ is a primitive root of p. For the primes < 211 these companion tables are complete for each p selected in connection with each irregular prime l. For the convenience of anyone who may wish to make use of these particular tables, I shall list values of $p > 1000$ that we employed in this connection.

1187, 1229, 1543, 1579, 1627, 1637, 1699, 1733, 1867, 2083, 2309, 2767, 2803, 3209, 3343, 3643, 4211, 4549, 4943, 5231, 5471, 5557, 9739.

All the numerical data obtained in the course of this work will ultimately be deposited in the library of the American Mathematical Society.

UNIVERSITY OF TEXAS.

NOTE

Since the above was written, Fermat's Last Theorem has been proved for the case $l = 587$. Hence, in view of Theorem 2, the theorem is true for all prime exponents less than 617.

ON THE NECESSARY CONDITIONS FOR THE MINIMUM OF A DOUBLE INTEGRAL

By Max Coral

1. **Introduction.** The purpose of the present paper is to exhibit an intimate connection which exists between the theory of the calculus of variations for a double integral and the corresponding theory for a simple integral. The variation problem which will be discussed is that of minimizing an integral

$$(1.1) \qquad \int\int_A f(x, y, z_1, \cdots, z_n, z_{1x}, \cdots, z_{nx}, z_{1y}, \cdots, z_{ny})\, dx\, dy$$

in a certain class of sets of functions $z_i(x, y)$ $(i = 1, 2, \cdots, n)$ all of which take on the same values on the boundary of the region A of integration. When it is not assumed that the minimizing set $Z_i(x, y)$ $(i = 1, 2, \cdots, n)$ have continuous partial derivatives of order greater than the first, the differential equations which must be satisfied by the minimizing set were first derived (at least for the case $n = 1$) by Haar [1],[1] [2], who made use of his so-called Fundamental Lemma of the calculus of variations for double integrals. A survey of the literature concerning this Fundamental Lemma will be found in the Chicago dissertation of Miss Huke [3]; the proof of the lemma was simplified considerably by Haar in his last paper [4] on the subject.

Of the further necessary conditions for a minimum of the integral (1.1), the analogue for double integrals of the Legendre condition for the minimum of a simple integral was first established by Mason [5] for the case $n = 1$. The analogue of the Weierstrass condition was first proved by E. E. Levi [6] for the case $n = 1$ and by McShane [7] for the general case. We shall not be concerned in this paper with the analogues of the Jacobi condition.

The Fundamental Lemma of Haar is unnecessary for the development of the theory of the calculus of variations for the integral (1.1) and the well-known Du Bois-Reymond lemma of the theory for simple integrals suffices. Indeed, it will be shown below that there is associated with the problem for the integral (1.1) an auxiliary minimum problem for a simple integral and that the necessary conditions of Haar, Weierstrass, and Legendre for the problem involving the integral (1.1) can be respectively deduced as simple corollaries of the necessary conditions of Euler, Weierstrass, and Legendre for the auxiliary problem. The condition of Haar will be derived below in a modified form involving integral equations.

Received May 24, 1937. The results contained in this paper were obtained in part by the author during his tenure of a National Research Fellowship.

[1] Numbers in square brackets refer to the bibliography at the end of the paper.

2. **Hypotheses.** Let A be a region of points (x, y) and let

$$z_i = Z_i(x, y) \qquad [(x, y) \text{ in } A;\ i = 1, 2, \cdots, n] \tag{2.1}$$

be functions which are continuous and which have continuous partial derivatives of the first order in A. The function $f(x, y, z, p, q)$ will be supposed to be defined and continuous for sets

$$(x, y, z, p, q) = (x, y, z_1, \cdots, z_n, p_1, \cdots, p_n, q_1, \cdots, q_n)$$

in a neighborhood G of the sets (x, y, Z, Z_x, Z_y) belonging to (2.1), and in G the function f shall have continuous derivatives of the first and second orders with respect to the arguments z_i, p_i, q_i $(i = 1, 2, \cdots, n)$.

We shall denote by $\mathfrak{M}$ the class of sets of functions

$$z_i = z_i(x, y) \qquad [(x, y) \text{ in } A;\ i = 1, 2, \cdots, n] \tag{2.2}$$

which have the following properties:

(a) the functions $z_i(x, y)$ are continuous in A and the region A may be decomposed into a finite number of subregions, each bounded by a simply closed regular curve, such that on each subregion the functions $z_i(x, y)$ have continuous partial derivatives of the first order;

(b) the elements (x, y, z, z_x, z_y) belonging to a set $z_i(x, y)$ are all in G;

(c) the functions $z_i(x, y)$ take on the same values on the boundary of A as do the functions of the set (2.1).

Under these hypotheses the integral

$$I = \int\int_A f(x, y, z, z_x, z_y)\, dx\, dy$$

has a finite real value when computed for any set $z_i(x, y)$ of the class $\mathfrak{M}$. We shall suppose that the set (2.1) furnishes a minimum to the integral I in the class $\mathfrak{M}$. The necessary conditions of Haar, Weierstrass, and Legendre which must then be satisfied by the minimizing set (2.1) will be derived below by consideration of an auxiliary minimum problem for a simple integral.

3. **The auxiliary minimum problem.** Consider any point (x_0, y_0) interior to A and let Γ represent a ring of points (x, y) defined by the equations

$$x = x_0 + \rho \cos \theta,\ y = y_0 + \rho \sin \theta \qquad (0 < r_1 \leqq \rho \leqq r;\ 0 \leqq \theta \leqq 2\pi), \tag{3.1}$$

where r is chosen so small that Γ lies interior to A. Let $T(\theta)$ be any function which is continuous on $0 \leqq \theta \leqq 2\pi$ and such that $T(0) = T(2\pi)$ and which has the further property that the interval $0 \leqq \theta \leqq 2\pi$ may be decomposed into a finite number of subintervals on each of which $T(\theta)$ has a continuous first derivative, and let $\mathfrak{R}$ represent the class of sets of functions $R_i(\rho)$ $(r_1 \leqq \rho \leqq r;$ $i = 1, 2, \cdots, n)$ which with their derivatives $R_i'(\rho)$ are continuous on $r_1 \leqq \rho \leqq r$ and have all their elements $[\rho, R(\rho), R'(\rho)]$ sufficiently near the sets

$(\rho, 0, 0)$ and for which $R_i(r_1) = R_i(r) = 0$ $(i = 1, 2, \cdots, n)$. For any set R_i of $\mathfrak{R}$ put

$$\zeta_i = Z_i(x_0 + \rho \cos \theta, y_0 + \rho \sin \theta) + R_i(\rho)T(\theta) \qquad (r_1 \leqq \rho \leqq r, 0 \leqq \theta \leqq 2\pi; i = 1, 2, \cdots, n). \tag{3.2}$$

Then in Γ we have

$$\begin{aligned} \frac{\partial}{\partial x}\zeta_i &= Z_{i_x} + R_i' T \cos \theta - R_i T' (\sin \theta/\rho), \\ \frac{\partial}{\partial y}\zeta_i &= Z_{i_y} + R_i' T \sin \theta + R_i T' (\cos \theta/\rho), \end{aligned} \qquad (i = 1, 2, \cdots, n). \tag{3.3}$$

If the definition of the functions ζ_i is extended so that $\zeta_i = Z_i$ in $A - \Gamma$ $(i = 1, 2, \cdots, n)$ the resulting set of functions clearly belongs to $\mathfrak{M}$. With $T(\theta)$ held fixed, $f(x_0 + \rho \cos \theta, y_0 + \rho \sin \theta, \zeta, \zeta_x, \zeta_y)$ becomes a function of (ρ, θ, R, R') as is clear from (3.2) and (3.3). If now we put

$$g(\rho, R, R') = \int_0^{2\pi} f(x_0 + \rho \cos \theta, y_0 + \rho \sin \theta, \zeta, \zeta_x, \zeta_y)\rho \, d\theta, \tag{3.4}$$

then $g(\rho, R, R')$ is defined for all R and R' sufficiently small and for all ρ such that the circle $x = x_0 + \rho \cos \theta$, $y = y_0 + \rho \sin \theta$ lies in A. Then the minimizing property of the set (2.1) implies that *the set* $R_i(\rho) \equiv 0$ $(r_1 \leqq \rho \leqq r;$ $i = 1, 2, \cdots, n)$ *furnishes a minimum to the integral*[2]

$$J = \int_{r_1}^{r} g[\rho, R(\rho), R'(\rho)]\, d\rho$$

in the class $\mathfrak{R}$.

4. The necessary condition of Haar. A first necessary condition which must be satisfied by the minimizing set (2.1) is contained in the following theorem:

THEOREM 1. *For every simply closed regular curve C lying interior to A the equations*

$$\int_C f_{p_i}\, dy - f_{q_i}\, dx = \int\int_C f_{z_i}\, dx\, dy \qquad (i = 1, 2, \cdots, n) \tag{4.1}$$

are satisfied, the arguments of f_{p_i}, f_{q_i} and f_{z_i} being the set (x, y, Z, Z_x, Z_y) belonging to the minimizing set (2.1).

The condition (4.1), while not in the form given by Haar, is equivalent to his system of differential equations. The theorem will be proved first for the case when C is a circle. Let C have its center at (x_0, y_0) and radius r. Let r_1 be any positive number less than r and consider the ring Γ with center at (x_0, y_0) and radii r_1 and r. For the corresponding auxiliary minimum problem, con-

[2] The integrand $g(\rho, R, R')$ possesses all the continuity properties needed to establish the first three necessary conditions for the auxiliary minimum problem. See Carathéodory, *Variationsrechnung und partielle Differentialgleichungen*, Leipzig, 1935, p. 190, footnote.

structed as described in the preceding section, the minimizing set $R_i(\rho) \equiv 0$ $(r_1 \leqq \rho \leqq r;\ i = 1, 2, \cdots, n)$ must satisfy the following equations, which are the Du Bois-Reymond form of the Euler equations for the problem:

$$g_{R_i'}(r, 0, 0) = \int_{r_1}^{r} g_{R_i}(\rho, 0, 0)\, d\rho + g_{R_i'}(r_1, 0, 0) \quad (i = 1, 2, \cdots, n)$$

or, by virtue of (3.4), with $T(\theta) \equiv 1$,

$$(4.2) \qquad \int_0^{2\pi} [(f_{p_i} \cos\theta + f_{q_i} \sin\theta)\rho]_{\rho=r_1}^{\rho=r}\, d\theta = \int_{r_1}^{r}\int_0^{2\pi} f_{z_i}\, \rho\, d\rho\, d\theta \quad (i = 1, 2, \cdots, n),$$

the arguments of $f_{p_i}, f_{q_i}, f_{z_i}$ in (4.2) being the set $[x, y, Z(x, y), Z_x(x, y), Z_y(x, y)]$ with x and y replaced by the values given in (3.1). Taking the limit as r_1 approaches zero in (4.2) one secures the desired equations (4.1) for the case of the circle C.

Now let C be any simply closed regular curve interior to A. Consider a closed region A' interior to A and containing C in its interior, and let $r > 0$ represent the minimum distance between the boundaries of A and A'. If $C_{xy,\rho}$ is a circle with center at (x, y) and radius ρ, then the functions

$$(4.3) \qquad \begin{aligned} M_{i\rho}(x, y) &= \int\int_{C_{xy,\rho}} f_{p_i}\, d\xi\, d\eta, \\ N_{i\rho}(x, y) &= \int\int_{C_{xy,\rho}} f_{q_i}\, d\xi\, d\eta, \\ P_{i\rho}(x, y) &= \int\int_{C_{xy,\rho}} f_{z_i}\, d\xi\, d\eta, \end{aligned} \qquad (i = 1, 2, \cdots, n)$$

in which the arguments of f_{p_i}, f_{q_i} and f_{z_i} are the set $[x + \xi, y + \eta, Z(x + \xi, y + \eta), Z_x(x + \xi, y + \eta), Z_y(x + \xi, y + \eta)]$, have the following properties:[3]

(a) for each value of ρ on the interval $0 < \rho \leqq r$ the functions (4.3) are single-valued and continuous for (x, y) in A';

(b) for each value of ρ on $0 < \rho \leqq r$ the functions (4.3) have continuous first partial derivatives with respect to x and y for (x, y) interior to A'; these derivatives are given by

$$(4.4) \qquad \begin{aligned} \frac{\partial}{\partial x} M_{i\rho}(x, y) &= \int_{C_{xy,\rho}} f_{p_i}\, d\eta, \\ \frac{\partial}{\partial y} M_{i\rho}(x, y) &= -\int_{C_{xy,\rho}} f_{p_i}\, d\xi, \end{aligned} \qquad (i = 1, 2, \cdots, n)$$

with similar formulas for the partial derivatives of $N_{i\rho}$ and $P_{i\rho}$;

[3] Property (c) is readily proved by means of the mean value theorem. For a proof of the analogue of property (b) for triple integrals see O. D. Kellogg, *Foundations of Potential Theory*, Berlin, 1929, p. 224.

(c) the functions (4.3) satisfy the conditions

$$\lim_{\rho\to 0} (1/\pi\rho^2)\, M_{i\rho}(x, y) = f_{p_i}[x, y, Z(x, y), Z_x(x, y), Z_y(x, y)],$$

$$\lim_{\rho\to 0} (1/\pi\rho^2)\, N_{i\rho}(x, y) = f_{q_i}[x, y, Z(x, y), Z_x(x, y), Z_y(x, y)], \quad (i = 1, 2, \cdots, n)$$

$$\lim_{\rho\to 0} (1/\pi\rho^2)\, P_{i\rho}(x, y) = f_{z_i}[x, y, Z(x, y), Z_x(x, y), Z_y(x, y)],$$

uniformly for (x, y) in A'.

Consider now any value ρ of the interval $0 < \rho \leqq r$. Since for each point (x, y) of C the equations (4.1) hold on the circle $C_{xy,\rho}$, we have

$$\begin{aligned}(1/\pi\rho^2)\int\int_C\left(\int\int_{C_{xy,\rho}} f_{z_i}\, d\xi\, d\eta\right)dx\, dy &= (1/\pi\rho^2)\int\int_C\left(\int_{C_{xy,\rho}} f_{p_i}\, d\eta - f_{q_i}\, d\xi\right)dx\, dy\\ &= (1/\pi\rho^2)\int\int_C\left[\frac{\partial}{\partial x}M_{i\rho}(x, y)\right.\\ &\qquad\left.+\frac{\partial}{\partial y}N_{i\rho}(x, y)\right]dx\, dy\\ &= (1/\pi\rho^2)\int_c M_{i\rho}(x, y)\, dy - N_{i\rho}(x, y)\, dx\\ &\qquad (i = 1, 2, \cdots, n),\end{aligned}$$

by virtue of Green's Theorem and property (b) above. Using property (c) we secure the equations (4.1) by taking the limit as ρ approaches zero.

5. The necessary condition of Weierstrass. The Weierstrass E-function is defined to be

$$\begin{aligned}E(x, y, z, p, q, P, Q) &= f(x, y, z, P, Q) - f(x, y, z, p, q)\\ &\quad - (P_i - p_i)f_{p_i}(x, y, z, p, q) - (Q_i - q_i)f_{q_i}(x, y, z, p, q),\end{aligned}$$

the repeated subscript i indicating summation with respect to i over the range $i = 1, 2, \cdots, n$. We have then the following result:

THEOREM 2. *At each point (x_0, y_0) of A the inequality*

$$E[x_0, y_0, Z(x_0, y_0), Z_x(x_0, y_0), Z_y(x_0, y_0), P, Q] \geqq 0 \tag{5.1}$$

holds for every set (P, Q) for which $[x_0, y_0, Z(x_0, y_0), P, Q]$ is in G and for which the matrix

$$\| P_i - Z_{ix}(x_0, y_0),\ Q_i - Z_{iy}(x_0, y_0) \| \tag{5.2}$$

has rank one.

Since the matrix (5.2) has rank one there exists a pair of constants $(a, b) \neq (0, 0)$ such that

$$a[P_i - Z_{ix}(x_0, y_0)] + b[Q_i - Z_{iy}(x_0, y_0)] = 0 \qquad (i = 1, 2, \cdots, n). \tag{5.3}$$

Let α $(0 \leqq \alpha \leqq 2\pi)$ be such that $\sin \alpha = a/(a^2 + b^2)^{\frac{1}{2}}$ and $\cos \alpha = -b/(a^2 + b^2)^{\frac{1}{2}}$ and determine constants R_i' $(i = 1, 2, \cdots, n)$ by the equations

$$(5.4)\qquad R_i' = [P_i - Z_{ix}(x_0, y_0)] \cos \alpha + [Q_i - Z_{iy}(x_0, y_0)] \sin \alpha \qquad (i = 1, 2, \cdots, n).$$

It follows from (5.3) and (5.4) that

$$\begin{aligned} R_i' \cos \alpha &= P_i - Z_{ix}(x_0, y_0), \\ R_i' \sin \alpha &= Q_i - Z_{iy}(x_0, y_0), \end{aligned} \qquad (i = 1, 2, \cdots, n).$$

The desired inequality (5.1) now becomes

$$(5.5)\qquad E(x_0, y_0, Z_0, Z_{0x}, Z_{0y}, Z_{0x} + R' \cos \alpha, Z_{0y} + R' \sin \alpha) \geqq 0,$$

where we have put $Z_{i0} = Z_i(x_0, y_0)$, $Z_{i0x} = Z_{ix}(x_0, y_0)$, $Z_{i0y} = Z_{iy}(x_0, y_0)$ $(i = 1, 2, \cdots, n)$.

Now consider again the auxiliary minimum problem constructed in §3. The minimizing set $R_i(\rho) \equiv 0$ $(r_1 \leqq \rho \leqq r; i = 1, 2, \cdots, n)$ must satisfy the Weierstrass necessary condition for that problem. Hence for $\rho = r_1$,

$$(5.6)\qquad g(r_1, 0, R') - g(r_1, 0, 0) - R_i' g_{R_i'}(r_1, 0, 0) \geqq 0.$$

After division by r_1 and the use of (3.4) and (3.3) we secure, by letting r_1 approach zero,

$$(5.7)\qquad \int_0^{2\pi} E[x_0, y_0, Z_0, Z_{0x}, Z_{0y}, Z_{0x} + R'T(\theta) \cos \theta, Z_{0y} + R'T(\theta) \sin \theta]\, d\theta \geqq 0.$$

Suppose that the inequality (5.5) is false. Then there exists an interval $0 < \alpha_1 \leqq \theta \leqq \alpha_2 < 2\pi$ to which α is interior, such that

$$(5.8)\qquad E(x_0, y_0, Z_0, Z_{0x}, Z_{0y}, Z_{0x} + R' \cos \theta, Z_{0y} + R' \sin \theta) < 0 \qquad (\alpha_1 \leqq \theta \leqq \alpha_2).$$

(Here and in what follows the modifications which are necessary in case $\alpha = 0$ or $\alpha = 2\pi$ will be obvious to the reader.) Choose a positive number ϵ so small that $\alpha_1 - \epsilon$ and $\alpha_2 + \epsilon$ are interior to $0 \leqq \theta \leqq 2\pi$ and define $T_\epsilon(\theta)$ as follows:

$$(5.9)\qquad T_\epsilon(\theta) = \begin{cases} (\epsilon + \theta - \alpha_1)/\epsilon & (\alpha_1 - \epsilon \leqq \theta \leqq \alpha_1), \\ 1 & (\alpha_1 \leqq \theta \leqq \alpha_2), \\ (\epsilon - \theta + \alpha_2)/\epsilon & (\alpha_2 \leqq \theta \leqq \alpha_2 + \epsilon), \end{cases}$$

and $T_\epsilon(\theta) = 0$ elsewhere on $0 \leqq \theta \leqq 2\pi$. For this function $T_\epsilon(\theta)$ the integral in (5.7) becomes

$$(5.10)\qquad \int_{\alpha_1 - \epsilon}^{\alpha_2 + \epsilon} E[x_0, y_0, Z_0, Z_{0x}, Z_{0y}, Z_{0x} + R'T_\epsilon(\theta) \cos \theta, Z_{0y} + R'T_\epsilon(\theta) \sin \theta]\, d\theta,$$

and this integral may be broken up as follows:

$$\int_{\alpha_1-\epsilon}^{\alpha_2+\epsilon} E\,d\theta = \int_{\alpha_1-\epsilon}^{\alpha_1} E\,d\theta + \int_{\alpha_1}^{\alpha_2} E\,d\theta + \int_{\alpha_2}^{\alpha_2+\epsilon} E\,d\theta, \tag{5.11}$$

the arguments of E being the same as in (5.10).

The second term on the right of (5.11) is negative by virtue of (5.8) and dominates the right side of (5.11) for sufficiently small ϵ, since the first and third terms clearly approach zero with ϵ. For our present choice of $T(\theta)$, then, we have a contradiction with the inequality (5.7), which must hold for all functions $T(\theta)$ of the type described in §3. Hence the inequality (5.5) is true.

6. **The Legendre condition.** The necessary condition of Legendre for the minimizing set (2.1) can be secured by well known means out of the Weierstrass condition. We shall derive it here, however, from a consideration of the auxiliary minimum problem.

THEOREM 3. *At each point* (x_0, y_0) *of* A *the inequality*

$$f_{p_ip_k}A_iA_k + 2f_{p_iq_k}A_iB_k + f_{q_iq_k}B_iB_k \geqq 0 \tag{6.1}$$

holds for every set (A, B) *for which the matrix*

$$\| A_i, B_i \| \tag{6.2}$$

has rank one. In the left member of the inequality (6.1) *the arguments of the derivatives of* f *are the set* $[x_0, y_0, Z(x_0, y_0), Z_x(x_0, y_0), Z_y(x_0, y_0)]$ *belonging to the minimizing set* (2.1).

As in the previous proof, determine a pair of constants $(a, b) \neq (0, 0)$ as solutions of the equations

$$aA_i + bB_i = 0 \qquad (i = 1, 2, \cdots, n) \tag{6.3}$$

and define α by the equations

$$\sin\alpha = a/(a^2 + b^2)^{\frac{1}{2}}, \qquad \cos\alpha = -b/(a^2 + b^2)^{\frac{1}{2}}, \qquad (0 \leqq \alpha \leqq 2\pi).$$

Determine constants C_i $(i = 1, 2, \cdots, n)$ by the equations

$$C_i = A_i\cos\alpha + B_i\sin\alpha \qquad (i = 1, 2, \cdots, n). \tag{6.4}$$

Then in consequence of (6.3) and (6.4) we have

$$C_i\cos\alpha = A_i, \quad C_i\sin\alpha = B_i, \qquad (i = 1, 2, \cdots, n)$$

and the inequality (6.1) becomes

$$(f_{p_ip_k}\cos^2\alpha + 2f_{p_iq_k}\cos\alpha\sin\alpha + f_{q_iq_k}\sin^2\alpha)C_iC_k \geqq 0. \tag{6.5}$$

Suppose the inequality (6.5) is false. An argument similar to that in the preceding section would lead to the existence of an interval $\alpha_1 \leqq \theta \leqq \alpha_2$ interior

to $0 \leqq \theta \leqq 2\pi$ and to a function $T_\epsilon(\theta)$, defined as in (5.9) for ϵ sufficiently small, such that

$$(6.6) \quad \int_0^{2\pi} (f_{p_i p_k} \cos^2 \theta + 2 f_{p_i q_k} \cos \theta \sin \theta + f_{q_i q_k} \sin^2 \theta)\; T_\epsilon(\theta)^2\, C_i C_k\, d\theta < 0,$$

the arguments of the derivatives of f being the set

$$[x_0\, , y_0\, , Z(x_0\, , y_0),\, Z_x(x_0\, , y_0),\, Z_y(x_0\, , y_0)].$$

On the other hand, the Legendre condition for the auxiliary minimum problem, which must be satisfied along the minimizing set $R_i(\rho) \equiv 0$ $(r_1 \leqq \rho \leqq r;$ $i = 1, 2, \cdots, n)$, requires that

$$g_{R_i' R_k'}(r_1\, , 0, 0) C_i C_k \geqq 0.$$

If this inequality is expressed in terms of the derivatives of the function f by means of (3.4), and the result is divided by r_1, the inequality which results in the limit as r_1 approaches zero is found to be in contradiction with (6.6). Hence the supposition that (6.5) is false is incorrect.

Bibliography

1. A. Haar, *Über die Variation der Doppelintegrale*, Journal für die reine und angewandte Mathematik, vol. 149 (1919), p. 1.
2. A. Haar, *Über eine Verallgemeinerung des Du Bois-Reymondschen Lemmas*, Acta Litterarum ac Scientiarum, vol. 1 (1922), p. 33.
3. A. Hüke, *An historical and critical study of the fundamental lemma of the calculus of variations*, in *Contributions to the Calculus of Variations, 1930*, Chicago, 1931, pp. 45–160.
4. A. Haar, *Zur Variationsrechnung*, Abhandlungen aus dem Mathematischen Seminar des Hamburgischen Univ., vol. 8 (1930), p. 1.
5. M. Mason, *A necessary condition for an extremum of a double integral*, Bulletin of the American Mathematical Society, vol. 13 (1907), p. 293.
6. E. E. Levi, *Sulla necessità della condizione di Weierstrass per l'estremo degli integrali doppi*, R. Accademia dei Lincei, Atti, vol. 24 (1915), p. 353.
7. E. J. McShane, *On the necessary condition of Weierstrass in the multiple integral problem of the calculus of variations*, Annals of Mathematics, vol. 32 (1931), p. 578 and p. 723.

Wayne University.

CONCERNING APPELL SETS AND ASSOCIATED LINEAR FUNCTIONAL EQUATIONS

By I. M. Sheffer

Introduction. We have elsewhere considered the linear difference equation with constant coefficients[1]

$$\sum_{j=1}^{k} a_j y(x + \omega_j) = F(x), \tag{1}$$

from the point of view of a local solution. That is, assuming only that $F(x)$ is analytic about a point, we have shown that $y(x)$ exists satisfying (1) in the neighborhood of this point. Now (1) is a particular case of the general linear differential equation of infinite order[2]

$$L[y(x)] \equiv \sum_{n=0}^{\infty} a_n y_{(x)}^{(n)} = F(x), \tag{2}$$

where if we set

$$L(t) \sim \sum_{0}^{\infty} a_n t^n \tag{3}$$

and call $L(t)$ the *generating function* for the operator $L[y]$, then the generating function for (1) is

$$L(t) = \sum_{j=1}^{k} a_j e^{\omega_j t}.$$

This suggests the possibility of developing a local theory for equation (2), at least when $L(t)$ is suitably restricted.

The solubility of equation (2) is linked with the problem of expanding $F(x)$ in a series of Appell polynomials $\{P_n(x)\}$ generated by $L(t)$; i.e., where the sequence $\{P_n(x)\}$ is defined by

$$L(t)e^{tx} \sim \sum_{n=0}^{\infty} P_n(x)t^n. \tag{4}$$

We see this formally from the fact that

$$L\left[\frac{x^n}{n!}\right] = P_n(x), \tag{5}$$

Received June 1, 1937; presented to the American Mathematical Society, April, 1937.

[1] Sheffer, Transactions of the American Mathematical Society, vol. 39 (1936), pp. 345–379, and vol. 41 (1937), pp. 153–159.

[2] For an investigation of equation (2) from another point of view see H. T. Davis, American Journal of Mathematics, vol. 52 (1930), pp. 97–108.

and therefore if $F(x)$ has the expansion

$$F(x) = \sum c_n P_n(x), \tag{6}$$

a formal solution of (2) is given by

$$y(x) = \sum c_n \frac{x^n}{n!}. \tag{7}$$

We are thus led to examine those functions $L(t)$ whose Appell polynomials can be used to expand the general analytic function.

For this purpose the following classification of functions $L(t)$ seems to be significant:

(i) $L(t) \sim \sum a_n t^n$ has a zero radius of convergence.

(ii) $L(t) = \sum a_n t^n$ has a finite, non-zero radius of convergence.

(iii) $L(t)$ is an entire function, not of finite exponential type.[3]

(iv) $L(t)$ is of finite exponential type.

The functions of classes (i) to (iii) appear to be inadequate to expand the general analytic function; we accordingly restrict our attention to functions of class (iv). We shall not prove the inadequacy of classes[4] (i) to (iii); rather, we shall give some examples to suggest the truth of the statement.

Suppose $L(t) = 1/(t - a)$ $(a \neq 0)$, so that L is of class (ii). One finds from (4) that

$$P_n(x) = -\frac{1}{a}\left[\frac{x^n}{n!} + \frac{x^{n-1}}{(n-1)!\,a} + \frac{x^{n-2}}{(n-2)!\,a^2} + \cdots + \frac{1}{a^n}\right],$$

or:

$$P_n(x) = -\frac{e^{ax}}{a^{n+1}}\left[\frac{\dfrac{1}{0!} + \dfrac{(ax)}{1!} + \cdots + \dfrac{(ax)^n}{n!}}{e^{ax}}\right].$$

Now the expression in brackets can be written $1 - r_n(x)$, where

$$r_n(x) = e^{-ax}\frac{(ax)^{n+1}}{(n+1)!}\left[1 + \frac{(ax)}{n+2} + \frac{(ax)^2}{(n+2)(n+3)} + \cdots\right];$$

and,

$$|r_n(x)| \leqq 2e^{|a|R}(|a|R)^{n+1}/(n+1)! \quad (n \geqq N_R \text{ sufficiently large}),$$

where $|x| \leqq R$. Again,

$$\sum c_n P_n(x) = -e^{ax}\sum \frac{c_n}{a^{n+1}} + e^{ax}\sum \frac{r_n(x)}{a^{n+1}}.$$

[3] $L(t) = \sum a_n t^n$ is of *finite exp. type* ρ if $\limsup_{n=\infty} |n!\,a_n|^{1/n} = \rho$.

[4] It would be of interest to have classes (i) to (iii) investigated, in order to determine the set of functions possessing a convergent expansion in the Appell polynomials generated by members of these classes.

We see, therefore, that $\sum c_n P_n(x)$ converges for at least one value of x if and only if $\sum c_n/a^n$ converges; and if convergent for one x it is convergent for all x. That is, when the series converges the sum is an entire function. Even so, not all entire functions can be expanded in this manner. We are thus far from having an expansion for the general analytic function.

As a second example, choose $L(t) = e^{-t^2/4}$, which is of class (iii). The associated Appell set is the set of Hermite polynomials, and for these it is known that the expansion converges in a horizontal strip of which the real axis is the bisector. Here again the general analytic function possesses no (convergent) expansion.

The generating functions $L(t)$ of class (iv) come closer to our aim; and it is the Appell sets generated by this class that we study in the present paper. In Part I we determine the character of the regions of convergence of Appell expansions. They are bounded by simple closed convex curves, and their position depends on the singularities of a function related to $L(t)$. In Part II we obtain a solution of equation (2); not, indeed, for all analytic functions, but rather for all functions that have a radius of convergence exceeding a suitably chosen number (depending on $L(t)$). This then permits us to find a class of functions possessing a convergent Appell expansion.

Part I. Convergence regions for series of Appell polynomials of class (iv)

Let

$$L(t) = \sum_0^\infty a_n t^n \tag{1.1}$$

be an entire function of finite exponential type (see footnote in Introduction), and let $\{P_n(x)\}$ be the set of Appell polynomials generated by L:

$$e^{tx} L(t) = \sum_0^\infty P_n(x) t^n. \tag{1.2}$$

Definition. Let $h(t) = \sum c_n t^n$ be of exp. type $\sigma < \infty$. Then the series $H(t) = \sum n! c_n t^n$ has a radius of convergence $1/\sigma$, and we say that *$h(t)$ is the Borel entire function associated with $H(t)$*. We shall denote this relation by $h(t) = \text{BEF}\ \{H(t)\}$.

Lemma 1.1. *For $n \geqq 1$,*

$$t^{n-1} e^{tx} = \text{BEF} \left\{ \frac{(n-1)! t^{n-1}}{(1 - tx)^n} \right\}. \tag{1.3}$$

The proof is straightforward, and need not be set down here.

Now $e^{tx} L(t) = \sum_0^\infty a_n (e^{tx} t^n)$, so that formally we have

$$\sum_0^\infty P_n(x) t^n = \text{BEF} \left\{ \sum_{n=0}^\infty a_n \cdot \frac{n! t^n}{(1 - tx)^{n+1}} \right\}. \tag{1.4}$$

Relation (1.4) is however more than formal; it is valid for x in any bounded region R and t sufficiently near to the origin (how near depending on R). For the expression in braces is uniformly convergent in x and t for x in R and $|t|$ sufficiently small, and may therefore be expanded in a convergent power series in t by Weierstrass' theorem. On doing so one finds for the coefficient of t^n the expression

$$a_0 x^n + na_1 x^{n-1} + n(n-1)a_2 x^{n-2} + \cdots + n! a_n,$$

so that (1.4) holds, since from (1.2) we obtain

$$P_n(x) = a_0 \frac{x^n}{n!} + a_1 \frac{x^{n-1}}{(n-1)!} + \cdots + a_n. \tag{1.5}$$

If in (1.4) we replace t by $1/t$ and divide through by t, we obtain

THEOREM 1.1. *Let $L(t)$ of* (1.1) *be of exponential type $\sigma < \infty$. Then*

$$\sum_{n=0}^{\infty} a_n \cdot \frac{n!}{(t-x)^{n+1}} = \sum_{n=0}^{\infty} \frac{n! P_n(x)}{t^{n+1}} \tag{1.6}$$

whenever both series converge. The right-hand series converges uniformly in x and t for x in any bounded region R and t in[5] *$|t| \geqq \rho > \sigma + d$, where d is the maximum value (or least upper bound) of $|x|$ in R. The left-hand series converges uniformly for x and t such that $|t - x| \geqq \tau > \sigma$. For x in any bounded region,* (1.6) *holds for all $|t|$ sufficiently large.*

Consider now the function $L^*(t)$ of which the BEF is $L(t)$:

$$L^*(t) = \sum_0^{\infty} n! a_n t^n. \tag{1.7}$$

Then

$$\frac{1}{t} L^*\left(\frac{1}{t}\right) = \sum \frac{n! a_n}{t^{n+1}}, \tag{1.8}$$

and on comparing with (1.6),

$$\frac{1}{t-x} L^*\left(\frac{1}{t-x}\right) = \sum \frac{n! P_n(x)}{t^{n+1}}, \tag{1.9}$$

valid for x in any bounded region and $|t|$ sufficiently large.

Equation (1.9) is an important relation, for by means of it we can acquire information concerning the behavior of $|n! P_n(x)|^{1/n}$ for large n. This is a consequence of the simple observation that the circle of convergence for the right-hand member of (1.9) is determined by the most distant singularity, of the left-hand member, from the origin.

Define $A(u)$ by

$$A(u) = uL^*(u), \tag{1.10}$$

[5] $e^{tx}L(t)$ is of exp. type $\leqq \sigma + |x|$.

and suppose for the moment that $A(u)$ is single-valued throughout its domain of existence. Let $G = \{\alpha\}$ be the set of all the singularities (including ∞ if necessary) of $A(u)$, and $E = \{\beta\}$ be the set of points defined by

$$\beta = -1/\alpha \tag{1.11}$$

as α runs through G. To each α (in G) corresponds a singularity $t = x + \alpha^{-1} = x - \beta$ of (1.9), and conversely. If then we define $D(x)$ by[6]

$$D(x) = \max \left| x + \frac{1}{\alpha} \right| = \max | x - \beta |, \tag{1.12}$$

we have

THEOREM 1.2. *For every x,*

$$\limsup_{n=\infty} | n!P_n(x) |^{1/n} = D(x). \tag{1.13}$$

This theorem is also true when $A(u)$ is a branch of a multiple-valued function, but in this case it is necessary to make precise those singularities of the complete analytic function "$A(u)$" (as we shall denote it) that are to be regarded as singularities of the particular branch $A(u)$ given by

$$A(u) = \sum n!a_n u^{n+1}. \tag{1.14}$$

From (1.4) we have

$$A\left(\frac{t}{1 - tx}\right) = \sum n!P_n(x)t^{n+1}, \tag{1.15}$$

valid for x in any bounded region on taking $| t |$ sufficiently small. *Let $x = \delta e^{i\theta} = \omega + i\eta$ be fixed*, and consider the transformation

$$u = \frac{t}{1 - tx}, \qquad t = \frac{u}{1 + ux}. \tag{1.16}$$

When t describes the circle $| t | = r$, u describes a circle C_r with the following radius and center:

$$R = r \,| r^2\delta^2 - 1 |^{-1}; \qquad (-\omega r^2/(r^2\delta^2 - 1),\ \eta r^2/(r^2\delta^2 - 1)).$$

Let B be the point $-\omega + i\eta$, and let l be the ray issuing from O (the origin) through B, and l' the ray in the opposite direction. Three cases are to be distinguished:

(i) $r < 1/\delta$: the center of C_r is on ray l'.

(ii) $r > 1/\delta$: the center of C_r is on ray l.

(iii) $r = 1/\delta$: C_r is a straight line (hereafter denoted by L_r) cutting l at right angles at a distance $1/2\delta$ from O. The distance from the center of C_r to O is $r^2\delta\, | r^2\delta^2 - 1 |^{-1}$.

[6] Point α runs over G and β over E. The set E is clearly bounded and closed, so that $\max | x - \beta |$ exists.

Let I_r denote the region in the u-plane into which $|t| < r$ maps under the transformation (1.16). It is then found that, according as we have cases (i), (ii) or (iii), I_r is respectively the interior of C_r, the exterior of C_r, the half-plane (determined by L_r) that contains ray l'. In every case the origin (O) is in I_r.

In case (i), R regarded as a function of r is an increasing function, varying from 0 at $r = 0$ to ∞ as $r \to 1/\delta$. On the other hand, for case (ii), the radius decreases to 0 as $r \to \infty$, the limiting center being the point H: $(-\omega + i\eta)/\delta^2$ (on l).

Let r increase from 0 to ∞. From 0 to $1/\delta$ the transformed curves are circles, C_{r_1} inside C_{r_2} for $r_1 < r_2$, all lying on the origin side of L_r, and filling out this half-plane. For $r = 1/\delta$ we get the line L_r. Then for r from $1/\delta$ to ∞ the circles form a decreasing set, C_{r_1} surrounding C_{r_2} for $r_1 < r_2$, all lying on the other side of L_r, and shrinking down to the point H as $r \to \infty$; and they fill out this other half-plane.

Now consider the complete analytic function "$A(u)$". For the given fixed x there is a smallest $r = r_1$ (> 0) such that C_{r_1} passes through a singularity[7] "α" of "$A(u)$"; (i.e., at least one branch of "A" is singular at "α"). We are to decide whether or not "α" is to be regarded as a singularity of $A(u)$ *relative*[8] *to the fixed* x. What we mean by this phrase will appear shortly. We consider three cases.

(i) $r_1 < 1/\delta$. It is clear that series (1.14) can be continued analytically throughout I_{r_1}, so that branch $A(u)$ is single-valued and analytic in I_{r_1}. If it is possible to continue this branch beyond C_{r_1} across "α"; i.e., if there is a circle containing "α" and therefore overlapping with I_{r_1}, such that series (1.14) can be continued from I_{r_1} into this circle; then "α" is *not* to be regarded as a singularity of $A(u)$ (relative to x). We shall say that the singularity "α" of "A" is *passed over*. If there is at least one singularity "α" on C_{r_1} that cannot be passed over, such a point is a singularity of $A(u)$ (relative to x), and the radius of convergence of (1.15) is then precisely r_1. If all singularities "α" on C_{r_1} are passed over, there will nevertheless exist a smallest value $r = r_2$ (> 0) such that on C_{r_2} there is at least one "α" that cannot be passed over. Such an "α" is a singularity of the branch $A(u)$ (relative to x), and series (1.15) has the radius of convergence r_2.

(ii) $r_1 > 1/\delta$. Here series (1.14) can be continued from O throughout the whole exterior of C_{r_1} (including ∞), and in this region I_{r_1} this branch is single-valued and analytic. The argument now follows the lines of case (i).

(iii) $r_1 = 1/\delta$. C_{r_1} is now the line L_{r_1}. If "α" is a finite singularity (of "A") on L_{r_1}, the method of case (i) tells us whether or not "α" is a singularity of branch $A(u)$ (relative to x). If there is no finite "α" on L_{r_1}, or if every such finite "α" is passed over, there is yet the possibility that "α" $= \infty$. (If ∞ is a

[7] We can ignore those singularities "α" of "A" that lie interior to the circle of convergence of (1.14), since these cannot be singularities of $A(u)$. Hence $r_1 > 0$.

[8] We shall presently see that if "α" is regarded as a singularity relative to one x then it can be regarded as a singular point relative to all x; i.e., it is independent of x.

singularity of "A" we consider it to lie on L_{r_1}.) It may not be possible to determine directly if branch $A(u)$ can be continued across "∞". We then proceed as follows:

If there exists an $r > 1/\delta$ such that $A(u)$ is single-valued and analytic in I_r (and therefore at ∞), then "∞" is passed over. But if for every $r > 1/\delta$, $A(u)$ does not remain single-valued and analytic in I_r, then ∞ is a singularity of $A(u)$ (relative to x).

To sum up: *For each fixed x* there is as described above a smallest positive number $r = r_x$ such that the branch $A(u)$ given by (1.14) can be continued analytically (and single-valuedly) throughout I_{r_x} but not throughout I_r for any $r > r_x$. On C_{r_x} there is at least one singularity "α" of "A" that is a singularity of $A(u)$ (relative to x). The radius of convergence of series (1.15) is r_x.

Let G_x be the set of those singularities of "A" on C_{r_x} that are not passed over; i.e., that are to be regarded as singularities of $A(u)$ (rel. to x). And let H_x be the set of those "α's" lying in and on C_{r_x} that are passed over, and E_x the set related to G_x by the transformation (1.11).

As x varies these sets can vary in their membership. Denote by G and E the respective logical sums of all the sets G_x and all the sets E_x. It is then seen that *Theorem* 1.2 *is valid with this choice of G and E* provided we show that a point "α" which is a singularity of $A(u)$ relative to an $x = x_1$ can never be a passed-over point relative to some other $x = x_2$; i.e., that the logical product $G_{x_1} \cdot H_{x_2}$ is[9] null for all x_1 and x_2.

Suppose it were possible to have an "α" so that $G_{x_1} \cdot H_{x_2} \neq 0$, and let $C_{r_1}^{(1)}$, $C_{r_2}^{(2)}$ be the circles corresponding to x_1 and x_2 in the description above. Then "α" is interior to $I_{r_2}^{(2)}$ but is on $C_{r_1}^{(1)}$. The closed segment joining 0 to "α" lies wholly in $I_{r_2}^{(2)}$, and (except for "α") wholly in $I_{r_1}^{(1)}$. If we continue $A(u)$ (as given by (1.14)) from 0 along this segment, we can pass over "α" (relative to x_2). But this same continuation obviously applies relative to x_1, so that "α" is passed over relative to x_1. Hence "α" is not in G_{x_1}, and this contradiction proves that (1.13) remains valid.

We shall call the locus $D(x) = c$ a *level curve* for the sequence $\{P_n(x)\}$. Concerning these level curves (to be denoted by J_c) we have the following properties.

(i) *$D(x)$ is a continuous function of x (for all finite x).*

This follows from the fact that $D(x)$ is a distance function, the set E being bounded and closed. Now $D(x) \to \infty$ as $|x| \to \infty$; hence $D(x)$ has *a minimum value*[10] D_m. We shall see presently that the value D_m is taken on at only one point.

[9] This amounts to saying that the set of singularities of $A(u)$ relative to a given x is independent of x in the sense that no two x's will yield contradictory information. According to our definition of G, we can regard G as the absolute set of singularities of the branch $A(u)$ given by (1.14). This is so even though there may be singularities "α" of "A" which never present themselves for test (by our above method) as to their *passed-over* or *singular* character relative to an x.

[10] If E consists of a single point, $D_m = 0$. In all other cases, $D_m > 0$.

From what has already been said, J_c ($c \geqq D_m$) is a bounded, closed set.

(ii) *For $c > D_m$, J_c consists of more than one point.*

If not, suppose J_c is the point x_0. Let x_1 be a point on J_{D_m} and x_2 on J_a where $a > c$. (Clearly, J_a is not-null for every $a \geqq D_m$.) Then an arc joining x_1 to x_2, and not passing through x_0, must contain at least one point of J_c. This provides a contradiction.

(iii) *For $c > D_m$, J_c has no isolated points.*

If otherwise, let x_0 on J_c be isolated. Then there is a neighborhood of x_0 (e.g., a sufficiently small circle K) such that for no point $x \neq x_0$ in K is $D(x) = c$. Therefore either $D(x) < c$ for all $x \neq x_0$ in K or $D(x) > c$ for all $x \neq x_0$ in K. Suppose $D(x) < c$. Let β_0, in E, be such that $d(x_0, \beta_0) \equiv |x_0 - \beta_0| = D(x_0)$. If the segment joining β_0 to x_0 be extended beyond x_0, then for x on this extension (and in K) we have $D(x) \geqq d(x, \beta_0) > D(x_0) = c$; a contradiction.

Now assume it possible to have $D(x) > c$. By (ii) there is a second point x_0' on J_c, and we assume x_0' to be chosen the closest of all points of $J_c - x_0$ to x_0. ($J_c - x_0$ is a closed set.) Let C, C' be circles of radius c and centers x_0, x_0'. All points of E lie in or on C and in or on C', and therefore in the closed zone Z common to C and C'. But for every x on the open segment joining x_0 to x_0', the circle C_x, with center at x and passing through the intersections of C and C', contains all of Z in or on it; and its radius is easily found to be less than c, so that $D(x) < c$. Since x can be chosen to lie in K, we thereby arrive at a contradiction.

(iv) *If x_1, x_2 are any points on J_c ($c \geqq D_m$), then every point x of the (open) segment joining x_1 to x_2 satisfies the relation $D(x) < c$.*

The proof follows from the latter part of the proof of (iii). Since $c = D_m$ is not excluded we arrive at the result:

(v) *The locus J_{D_m} consists of a single point (denoted by x^*).*

This is important when we come to consider expansions in $\{P_n(x)\}$-series, for it tells us that all regions of convergence of such series have the unique point x^* in common. Another corollary of (iv) is

(vi) *The locus J_c cannot contain three or more collinear points.*

Let l_θ be any *ray* (i.e., a *half*-line) issuing from x^* at the angle θ, $0 \leqq \theta \leqq 2\pi$.

(vii) *The ray l_θ meets each locus J_c ($c > D_m$) once and only once.*

First we observe that since $D(x) \to \infty$ as $|x| \to \infty$, every locus J_c ($c > D_m$) is met at least once by l_θ. Suppose l_θ meets J_c in two or more points, and let x_1, x_2 be two such. By (iv), if x_3 lies on l_θ between x_1 and x_2, then $D(x_3) = c' < c$. Then by continuity there is an x_4 between x^* and x_1 (supposing x_1 closer to x^* than is x_2) for which $D(x_4) = c'$; so that x_1, lying between x_4 and x_3, must satisfy the relation $D(x_1) < c' < c$, a contradiction.

Consider the locus J_c, $c > D_m$. If we introduce a polar coördinate system (r, θ), using x^* as the pole ($r = 0$), we have shown that for each θ there is one and only one r. That is, if we write the locus J_c as $r = F_c(\theta)$, then *$F_c(\theta)$ is a single-valued function of θ, $0 \leqq \theta \leqq 2\pi$.* We can however say more:

(viii) *The function $r = F_c(\theta)$ is continuous for all θ.*

For suppose $\theta = \theta_0$ is a point of discontinuity, and let $r_0 = F_c(\theta_0)$. Since $F_c(\theta)$ is a bounded function, there exists a sequence $\{\theta_n\} \to \theta_0$ such that $\{r_n = F_c(\theta_n)\}$ has a limit $r_0' \neq r_0$. Now $x_n = (r_n, \theta_n)$ is a point of J_c. Hence by continuity of $D(x)$, lim $(r_n, \theta_n) = (r_0', \theta_0) = x_0'$ is a point of J_c. Now x_0' is on ray l_{θ_0}, and $x_0' \neq x_0$. We thus contradict (vii), so the assumed discontinuity cannot exist.

Combining these results we can state the

THEOREM 1.3. *The locus J_{D_m} is a single point x^*. For every $c > D_m$, J_c is a simple closed convex curve, containing in its interior the point x^* and the locus $J_{c'}$ for every $c' < c$.*

If the set of singularities of $A(u)$ is finite in number,[11] the loci J_c consist of circular arcs, each J_c being made up of a finite number of such arcs, of radius c. There can be arcs of circles even in more complicated cases, but if $A(u)$ has, for example, curves of singularities or regions of singularities, then the loci may not have this simple character.

From Theorem 1.2 we immediately have

THEOREM 1.4. *Let* lim sup $| c_n/n! |^{1/n} = \rho < 1/D_m$ *and set* $c = 1/\rho$. *The series*

$$F(x) = \sum_0^\infty c_n P_n(x) \tag{1.17}$$

converges absolutely for every x interior to J_c, and converges uniformly in every closed region interior to J_c. The sum function $F(x)$ is analytic within J_c.

One naturally inquires if there can exist points of convergence beyond J_c. This is sometimes the case, and was indeed encountered in the treatment of equation (1) of the Introduction.[12]

Before going on to the solution of the linear differential equation (2) of the Introduction, we wish to examine the exponential type of the function $e^{tx}L(t)$. Let $L(t)$ be of exp. type ρ ($< \infty$), and let $e^{tx}L(t)$ be of exp. type $\tau = \tau(x)$. e^{tx} is itself of exp. type $| x |$. We have

LEMMA 1.2. *For all x,*

$$\tau(x) = D(x). \tag{1.18}$$

For the radius of convergence of a series $\sum b_n t^n$ is the reciprocal of the exp. type of its BEF, $\sum b_n t^n/n!$. Now $e^{tx}L(t) = \text{BEF}\{\sum n! P_n(x)t^n\}$, so that from (1.13), $\tau(x) = D(x)$.

From this follows

LEMMA 1.3. *For all x,*

$$| \rho - | x | | \leqq \tau = D(x) \leqq \rho + | x |. \tag{1.19}$$

For, the radius of convergence of $A(t) = \sum n! a_n t^{n+1}$ is $1/\rho$, so that the α point nearest the origin satisfies the relation $| \alpha | = 1/\rho$; therefore the β point

[11] This was the case in our Transactions paper (already cited).

[12] Loc. cit., Transactions, vol. 39 (1936), p. 355, footnote.

farthest from the origin is such that $|\beta| = \rho$. Since $D(x) = \max |x - \beta|$, we obtain $\tau(x) = D(x) \leqq |x| + \rho$. Again, the point β nearest to x satisfies the condition $|\beta - x| \geqq |\rho - |x||$, as a diagram readily shows. Hence $\tau \geqq |\beta - x| \geqq |\rho - |x||$.

Both limits in (1.19) can be attained. Consider, for example, $L(t) = e^{\rho t}$. Then $\tau(\rho) = 2\rho = \rho + |x|$, and $\tau(-\rho) = 0 = \rho - |x|$.

Part II. The associated linear functional equation: a semi-local solution

Let

$$L(t) = \sum_0^\infty a_n t^n \tag{2.1}$$

again be of finite exp. type ρ, so that

$$\limsup_{n=\infty} |n! a_n|^{1/n} = \rho < \infty. \tag{2.2}$$

We consider the linear differential equation of infinite order generated by L:

$$L[y(x)] \equiv \sum_0^\infty a_n y^{(n)}(x) = F(x). \tag{2.3}$$

It is no restriction to suppose[13] that $a_0 \neq 0$. Then $K(t) = 1/L(t)$ exists as a convergent power series in t, and if we define $\{Q_n(x)\}$ as the Appell set for $K(t)$, we have

$$K\left[\frac{x^n}{n!}\right] = Q_n(x), \tag{2.4}$$

or, what is equivalent (as is readily established),

$$L[Q_n(x)] = \frac{x^n}{n!}. \tag{2.5}$$

From the relation between Appell set and generating function (see (4) of Introduction) we also have

$$K(t)e^{tx} = \sum_0^\infty Q_n(x)t^n, \tag{2.6}$$

so that if C is a contour around $t = 0$, lying within the circle of convergence of $K(t)$, then

$$Q_n(x) = \frac{1}{2\pi i}\int_C \frac{e^{tx}}{L(t)t^{n+1}}\,dt. \tag{2.7}$$

The set $\{Q_n\}$ is less suitable for expansion purposes than is $\{P_n\}$ of Part I, since the generating function for $\{Q_n\}$ is only of class (ii). But by altering

[13] For suppose $a_i = 0$, $i = 0, 1, \cdots, p-1$, $a_p \neq 0$, and let $L_1(t) = L(t)/t^p = a_p + a_{p+1}t + \cdots$. If we can solve $L_1[y(x)] = G(x)$, then on differentiating p times we get $L[y(x)] = G^{(p)}(x)$, so that (2.3) is solved on choosing G to satisfy $G^{(p)}(x) = F(x)$.

the contour C we are led to a sequence of functions much more serviceable than $\{Q_n\}$. The method used is due initially to Hurwitz[14] (who was concerned with the simple difference equation $\Delta y(x) = F(x)$, $F(x)$ entire).

Since $K(t) = 1/L(t)$ is a meromorphic function the right-hand member of (2.7) will define a function no matter how distant the contour C is from the origin, provided that C does not pass through any singularity of $K(t)$. Choose such a contour $C \equiv C_n$, of radius r_n, center at the origin, and denote the corresponding function by Q_{n,r_n}:

$$Q_{n,r_n}(x) = \frac{1}{2\pi i}\int_{C_n} \frac{e^{tx}}{L(t)t^{n+1}}\,dt. \tag{2.8}$$

On applying operator L to both sides of (2.8), and taking L under the integral sign (an operation which is easily shown to be valid), we obtain

$$L[Q_{n,r_n}(x)] = \frac{x^n}{n!}. \tag{2.9}$$

We now examine the modulus of (2.8). From (2.1) and (2.2) we see that for every $\epsilon > 0$ there is an $N = N_\epsilon$ such that

$$M(r) = \max_{|t|=r} |L(t)| \leqq Ne^{(\rho+\epsilon)r}. \tag{2.10}$$

Again, there is the following theorem on entire functions:[15]

Let $f(z)$ be of finite order and let k be any number exceeding one. There is a number $H = H(k)$ such that for every $R > 0$ the closed interval (R, kR) contains at least one number r for which

$$|f(z)| > [M(kr)]^{-H}$$

for all z on $|z| = r$.

Here again $M(r)$ is the maximum modulus of $f(z)$ on $|z| = r$.

Set $k = 1 + \delta$, $\delta > 0$, and apply the theorem to $L(t)$. It gives

$$M(r) > [M(r(1 + \delta))]^{-H},$$

or

$$\left|\frac{1}{L(t)}\right| < [M(r(1 + \delta))]^H, \tag{2.11}$$

where, r' being any number, the number r exists somewhere in $r' \leqq r \leqq r'(1 + \delta)$. As we are going to vary our contours C_n, let us replace r', r by r_n', r_n. We take r_n (which depends on r_n') as the radius of C_n. Then from (2.8):

$$|Q_{n,r_n}(x)| \leqq \frac{1}{r_n^n}[M(r_n(1 + \delta))]^H \cdot e^{|x|r_n} \leqq r_n'^{-n} \cdot e^{|x|r_n'(1+\delta)} \cdot [M(r_n'(1 + \delta)^2)]^H;$$

and from (2.10):

$$|Q_{n,r_n}(x)| \leqq N \cdot r_n'^{-n} \cdot e^{r_n'(1+\delta)[|x|+H(\rho+\epsilon)(1+\delta)]}. \tag{2.12}$$

[14] Acta Mathematica, vol. 20 (1896–97), pp. 285–312.

[15] G. Valiron, *Lectures on the General Theory of Integral Functions*, Toulouse, 1923, p. 89.

Now r_n' is arbitrary. Let us choose

$$r_n' = dn, \tag{2.13}$$

d being independent of n, and make use of the asymptotic relation

$$n! \sim (2\pi n)^{\frac{1}{2}} \left(\frac{n}{e}\right)^n.$$

Then

$$n! \mid Q_{n,r_n}(x) \mid \leqq N \cdot \frac{(2\pi n)^{\frac{1}{2}}}{d^n} \cdot e^{n\{-1+d(1+\delta)[\mid x \mid + H(\rho+\epsilon)(1+\delta)]\}}, \tag{2.14}$$

and

$$\limsup \mid n! Q_{n,r_n}(x) \mid^{1/n} \leqq \frac{1}{d} \cdot e^{\{-1+d(1+\delta)[\mid x \mid + H(\rho+\epsilon)(1+\delta)]\}}. \tag{2.15}$$

Consider the right-hand member as a function of d. It is found that there is a minimum for $d = \{(1 + \delta)[\mid x \mid + H(\rho + \epsilon)(1 + \delta)]\}^{-1}$. In order to have d independent of x, we set $x = 0$. Furthermore, since $\epsilon > 0$ is arbitrary, it is suggested that we choose

$$d = \{H\rho(1 + \delta)^2\}^{-1}. \tag{2.16}$$

The right-hand member of (2.15) becomes

$$H\rho(1 + \delta)^2 \cdot e^{\{(\epsilon/\rho) + \mid x \mid [H\rho(1+\delta)]^{-1}\}}.$$

Now the left side of (2.15) is independent of ϵ. We therefore obtain[16]

$$\limsup \mid n! Q_{n,r_n}(x) \mid^{1/n} \leqq H\rho(1 + \delta)^2 \cdot e^{\mid x \mid / H\rho(1+\delta)}. \tag{2.17}$$

Let $\{c_n\}$ be a sequence with $\limsup \mid c_n \mid^{1/n} = \sigma$, so that

$$\limsup \mid n! c_n Q_{n,r_n}(x) \mid^{1/n} \leqq \sigma H\rho(1 + \delta)^2 \cdot e^{\mid x \mid / H\rho(1+\delta)}. \tag{2.18}$$

The right-hand side will be less than one for $\mid x \mid$ sufficiently small if $\sigma H\rho(1 + \delta)^2 < 1$. We can therefore state

THEOREM 2.1. *Let $L(t)$ be of exponential type ρ $(< \infty)$. There is a number $\Delta > 0$, depending only*[17] *on $L(t)$, such that if*

$$\limsup \mid c_n \mid^{1/n} = \sigma < \frac{1}{\rho\Delta}, \tag{2.19}$$

then series

$$\sum_{n=0}^{\infty} n! c_n Q_{n,r_n}(x) \tag{2.20}$$

converges absolutely and uniformly in some neighborhood of $x = 0$.

[16] Since $\delta > 0$ is arbitrary it might be thought that we can let $\delta \to 0$. But H depends on δ in a manner that is not specified in the theorem (in which H and $k = 1 + \delta$ first appear); we do not therefore know what value of δ will make the right-hand member of (2.17) a minimum.

[17] Δ can be taken as the minimum (or greatest lower bound) of $H(1 + \delta)^2$ for all $\delta > 0$.

COROLLARY. *There is a number $\omega > 0$, depending*[18] *only on $L(t)$ and σ, such that* (2.20) *converges absolutely for all $|x| < \omega$, and converges uniformly in every closed region in $|x| < \omega$.*

When the exponential type ρ is *zero*, there is a marked simplification in the above theorem. For consider (2.15) where now we have set $\rho = 0$. Since $\epsilon > 0$ is arbitrary, we can let $\epsilon \to 0$. This eliminates the term $H\epsilon(1 + \delta)$. Again, since $\delta > 0$ is arbitrary, we can let $\delta \to 0$. The left-hand member is independent of ϵ and δ, so we obtain

$$\text{(2.21)} \qquad \limsup |n!\,Q_{n,r_n}(x)|^{1/n} \leqq \frac{1}{d}\cdot e^{-1+d|x|},$$

$$\text{(2.22)} \qquad \limsup |n!\,c_n\,Q_{n,r_n}(x)|^{1/n} \leqq \frac{\sigma}{ed}\cdot e^{d|x|}.$$

Now *choose*[19] $d = \sigma$ (where for the moment we assume that $\sigma \neq 0$). Then

$$\text{(2.23)} \qquad \limsup |n!\,c_n Q_{n,r_n}(x)|^{1/n} \leqq e^{\sigma|x|-1}.$$

From this follows

THEOREM 2.2. *Let $L(t)$ be of exponential type zero. To*[20] *every $\sigma < \infty$ there is a sequence of contours C_n such that if* $\limsup |c_n|^{1/n} = \sigma$ *then series* (2.20) *converges absolutely for all x in $|x| < 1/\sigma$ and converges uniformly in every closed region in $|x| < 1/\sigma$.*

LEMMA 2.1. *Let $L(t)$ be of exponential type ρ $(< \infty)$. Then*

$$\text{(2.24)} \qquad L[y(x)] \equiv \sum_0^\infty a_n y^{(n)}(x)$$

converges absolutely in $|x - x_0| < \lambda - \rho$ if $y(x)$ is analytic in the circle $|x - x_0| < \lambda$, $(\lambda > \rho)$; and the convergence is uniform in every closed region in $|x - x_0| < \lambda - \rho$.

To show this, let C be a circle with center at x_0 and radius λ', where $\rho < \lambda' < \lambda$. For x in $|x - x_0| < \lambda' - \rho$ we have

$$y^{(n)}(x) = \frac{n!}{2\pi i}\int_C \frac{y(t)\,dt}{(t-x)^{n+1}},$$

[18] If δ', H' are the values giving rise to Δ (see previous footnote), then ω is defined by

$$\omega = -\rho H'(1 + \delta') \log [\sigma\rho\Delta].$$

This is seen by finding for what values of x the right-hand member of (2.18) is less than unity. If we set

$$A = -\rho H'(1 + \delta'), \qquad B = A \log (\rho\Delta),$$

then we can write

$$\omega = B + A \log \sigma,$$

where A and B depend only on $L(t)$, not on σ.

[19] This means that the contours C_n, which serve to define the functions $Q_{n,r_n}(x)$, change with σ, so that for different σ's we may be dealing with different sequences $\{Q_{n,r_n}(x)\}$.

[20] The preceding relation $d = \sigma$ which furnishes the proof requires that $\sigma \neq 0$, but Theorem 2.2 is readily seen to hold even if $\sigma = 0$.

and therefore

$$L[y] = \frac{1}{2\pi i}\int_C y(t)\left\{\sum_{n=0}^{\infty}\frac{n!\,a_n}{(t-x)^{n+1}}\right\}dt,$$

since the series converges uniformly (and absolutely) for t on C and x in any closed region in $|x - x_0| < \lambda' - \rho$. On letting λ' increase toward λ we obtain the desired conclusion.

COROLLARY. *If $\rho = 0$, then $L[y(x)]$ converges in the neighborhood of every point x_0 at which $y(x)$ is analytic, and the convergence reaches out to the circle of convergence of the power series of $y(x)$ about x_0.*

LEMMA 2.2. *With the hypotheses of Lemma 2.1, if*

$$y(x) = \sum_0^{\infty} y_n(x - x_0)^n, \tag{2.25}$$

the radius of convergence being λ, then

$$L[y(x)] = \sum_{n=0}^{\infty} y_n L[(x - x_0)^n], \tag{2.26}$$

the latter series converging uniformly in every closed region in $|x - x_0| < \lambda - \rho$.

Since $\limsup |y_n|^{1/n} = 1/\lambda$, to every $\epsilon > 0$ corresponds an $a = a_\epsilon$ such that

$$|y_n| < a\left(\frac{1}{\lambda} + \epsilon\right)^n;$$

hence

$$y(x) \ll a\sum\left[\left(\frac{1}{\lambda} + \epsilon\right)|x - x_0|\right]^n = \frac{a}{1 - \left(\frac{1}{\lambda} + \epsilon\right)|x - x_0|}.$$

Therefore on setting $u = (1/\lambda + \epsilon)\,|x - x_0|$,

$$L[y] \ll a\sum_{n=0}^{\infty}\frac{n!\,|a_n|\left(\frac{1}{\lambda} + \epsilon\right)^n}{(1 - u)^{n+1}}. \tag{2.27}$$

This is valid provided that $|u| < 1$ (a condition which is true if x lies in any closed region in $|x - x_0| < \lambda$, when ϵ is chosen sufficiently small), and that $|1 - u| > \rho(1/\lambda + \epsilon)$. This latter condition is fulfilled if

$$|x - x_0| < \frac{\lambda - \rho - \rho\lambda\epsilon}{1 + \lambda\epsilon},$$

and from the arbitrariness of ϵ we need only have $|x - x_0| < \lambda - \rho$. Absolute and uniform convergence of (2.27) in any closed region in $|x - x_0| < \lambda - \rho$ is immediate. From the absolute convergence it follows that the double series

$$L[y] = L\left[\sum_0^{\infty} y_n(x - x_0)^n\right] = \sum_{n=0}^{\infty} a_n\left\{\sum_{k=0}^{\infty} y_k(x - x_0)^k\right\}^{(n)}$$

converges absolutely, and may therefore be summed in any way. But one way of summing is to take L under the Σ-sign: $\sum_0^\infty y_n L[(x - x_0)^n]$. This establishes the lemma.

COROLLARY. *If $\rho = 0$, then (2.26) is valid in $|x - x_0| < \lambda$.*

LEMMA 2.3. *Let $\{u_n(x) = \sum_{k=0}^\infty b_{nk}(x - x_0)^k\}$, $n = 0, 1, \cdots$, be analytic in $|x - x_0| < \lambda$, $\lambda > \rho$, and let the double series $\sum_{n,k=0}^\infty |b_{nk}(x - x_0)^k|$ converge for $|x - x_0| < \lambda$. Then*

$$L\left[\sum_{n=0}^\infty u_n(x)\right] = \sum_{n=0}^\infty L[u_n(x)], \tag{2.28}$$

both series being uniformly convergent for x in any closed region in $|x - x_0| < \lambda - \rho$.

The proof is patterned after that of Lemma 2.2, and need not be given in detail. Essentially it consists in showing that the double series

$$\sum_{n=0}^\infty a_n\{u_0^{(n)} + u_1^{(n)} + u_2^{(n)} + \cdots\}$$

is absolutely convergent in $|x - x_0| < \lambda - \rho$ (uniformly in any closed region therein), and may therefore be summed in any way. One such summing yields the right-hand member of (2.28).

Now the series (2.20), with x replaced by $x - x_0$, fulfills the conditions of Lemma 2.3 when σ is so small that $\omega > \rho$. For, (2.8) gives us the series

$$Q_{n,r_n}(x - x_0) = \sum_{k=0}^\infty \frac{(x - x_0)^k}{k!}\left(\frac{1}{2\pi i}\int_{c_n} \frac{t^{k-n-1}}{L(t)}\,dt\right),$$

convergent for all x. We must therefore show that series

$$\sum_{n,k=0}^\infty \frac{|x - x_0|^k}{k!}\cdot\left|\frac{n!\,c_n}{2\pi i}\int_{c_n}\frac{t^{k-n-1}}{L(t)}\,dt\right| \tag{a}$$

converges for $|x - x_0| < \omega$. On using inequalities (2.10) and (2.11) in the manner already considered, and choosing $r_n' = dn$, series (a) is found to be less than

$$N'\sum_{n=0}^\infty \frac{n^{\frac{1}{2}}\,|c_n|}{(ed)^n}\cdot e^{nd(1+\delta)\{H(\rho+\epsilon)(1+\delta)+|x-x_0|\}}. \tag{b}$$

We can let $\epsilon \to 0$. On again choosing $d = \{H\rho(1 + \delta)^2\}^{-1}$, we see that (b) converges provided x is such that $\sigma H\rho(1 + \delta)^2\cdot e^{|x-x_0|\{H\rho(1+\delta)\}^{-1}} < 1$. Since $\Delta = H(1 + \delta)^2$, it now follows that (a) converges for all x in $|x - x_0| < \omega$, as was to be shown.

This gives us the relation

$$L\left[\sum_0^\infty n!c_nQ_{n,r_n}(x-x_0)\right]=\sum_0^\infty n!c_nL[Q_{n,r_n}(x-x_0)];$$

and on using (2.9) (with x replaced by $x-x_0$), we have

THEOREM 2.3. *Let $L(t)$ be of exponential type ρ ($<\infty$). Let* $\limsup|c_n|^{1/n}=\sigma$ *be so small that*[21] $\omega>\rho$. *Then*

$$L\left[\sum_{n=0}^\infty n!c_nQ_{n,r_n}(x-x_0)\right]=\sum_{n=0}^\infty c_n(x-x_0)^n, \tag{2.29}$$

the left side being convergent in $|x-x_0|<\omega-\rho$ and the right side convergent in $|x-x_0|<1/\sigma$.

Otherwise stated, Theorem 2.3 gives us a *semi-local* existence theorem for the equation (2.3):

THEOREM 2.4. *Let $L(t)$ be of exponential type ρ ($<\infty$). If $F(x)$ is analytic about $x=x_0$ in a circle of radius $r>1/\sigma^*$, then equation* (2.3) *possesses an analytic solution $y(x)$ in the neighborhood of $x=x_0$. Moreover, if*

$$F(x)=\sum_0^\infty c_n(x-x_0)^n, \tag{2.30}$$

then $y(x)$ has the form

$$y(x)=\sum_0^\infty n!c_nQ_{n,r_n}(x-x_0), \tag{2.31}$$

this latter series converging in $|x-x_0|<\omega-\rho$, where ω is determined for the value[22] $\sigma=1/r$.

We refer to this as a *semi-local* solution because our method does not permit the circle of analyticity of $F(x)$ about $x=x_0$ to become smaller than $1/\sigma^*$. There is however no reason to believe that this restriction, which results from the method used, is inherent in the problem. The question remains open.

If $\rho=0$ this restriction automatically disappears, so that we have

THEOREM 2.5. *Let $L(t)$ be of exponential type zero. To every function $F(x)$ analytic about $x=x_0$ there is a function $y(x)$, also analytic about $x=x_0$, which satisfies equation* (2.3); *and if* (2.30) *has the radius of convergence r, then* (2.31) *also converges*[23] *for $|x-x_0|<r$.*

These solutions, semi-local or local as the case may be, permit us to return to the problem of Appell expansions. Let $F(x)$ be analytic about $x=0$ in a

[21] We saw that $\omega=B+A\log\sigma$ where A, B are independent of σ and $A<0$. Hence $\lim\omega=+\infty$ as $\sigma\to+0$, so that values of σ exist for which $\omega>\rho$. Let us denote by σ^* that value such that $\sigma<\sigma^*$ implies $\omega>\rho$. Then

$$\sigma^*=e^{(\rho-B)/A}.$$

[22] In particular, if $\sigma=0$ then $\omega=\infty$, so that *equation* (2.3) *possesses an entire function solution when $F(x)$ is entire.*

[23] In particular, if $F(x)$ is an entire function, so is $y(x)$.

circle of radius $r > 1/\sigma^*$, so that Theorem 2.4 holds, and let $y(x) = \sum y_n x^n$ be the solution. Its radius of convergence is at least $\lambda = \omega - \rho$. Suppose $\lambda > \rho$, so that $\omega > 2\rho$. Then by Lemma 2.2,

$$L[y(x)] = \sum_0^\infty y_n L[x^n] = \sum_0^\infty n!\, y_n P_n(x),$$

where $\{P_n(x)\}$ is the Appell set generated by $L(t)$. But $L[y(x)] = F(x)$. Accordingly we have

THEOREM 2.6. *Let $L(t)$ be of exponential type ρ ($< \infty$) and $\{P_n(x)\}$ the corresponding Appell set. If $F(x)$ is analytic about $x = 0$ in a circle of radius $r = 1/\sigma$, where σ is so small that $\omega > 2\rho$, then $F(x)$ possesses the Appell expansion*

$$F(x) = \sum_{n=0}^\infty n!\, y_n P_n(x), \tag{2.32}$$

valid in the neighborhood of $x = 0$. Here $\{y_n\}$ is the set of coefficients of the solution $y(x) = \sum y_n x^n$ of (2.3) given by Theorem 2.4.

This theorem is not as satisfactory as one might wish. It refers $F(x)$ to the origin, whereas we know from Part I that the central point in P_n-expansions is the point x^*. The difficulty resides in the semi-local character of the solution of (2.3).

When $\rho = 0$ this difficulty disappears. For then $A(t)$ is an entire function. Its only singularity is $\alpha = \infty$, so that $\beta = 0$. Hence $x^* = 0$, and the level curves are concentric circles with center at the origin. Application of Theorem 2.5 and Lemma 2.2 gives

THEOREM 2.7. *Let $L(t)$ be of exponential type zero. Then every function $F(x)$ analytic about the origin possesses a P_n-expansion that is valid in the same circle as is the power series for $F(x)$.*

This shows that the Appell expansions corresponding to a function $L(t)$ of exp. type zero are markedly like ordinary power series in their convergence properties.

PENNSYLVANIA STATE COLLEGE.

UNIFORMITY PROPERTIES IN TOPOLOGICAL SPACE SATISFYING THE FIRST DENUMERABILITY POSTULATE

By L. W. Cohen

Theorems involving convergence, completeness and uniform continuity are consequences of what may be called the uniformity properties which inhere in metric spaces. It is perhaps of some interest to seek in spaces which are not metrizable similarly effective uniformity properties, in terms of which the theorems referred to may be reformulated.[1] It is the purpose of this note to do this. The spaces considered are the topological spaces of Hausdorff satisfying the first denumerability postulate, namely, that to each $p \subset S$ the topology at p is determined by a sequence of neighborhoods $U_n(p)$. The uniformizing entity is the class $[U_n]$ of all $U_n(p)$ for fixed n and all $p \subset S$. We make the following definitions:

1. A sequence $p_k \subset S$ is a *Cauchy sequence* if for each n there is a k_n and a $q_n \subset S$ such that $p_k \subset U_n(q_n)$ if $k > k_n$.
2. A space S is *complete* if every Cauchy sequence has a limit.
3. A set $M \subset S$ is *totally bounded* if for each n there are points

$$p_{n,1}, p_{n,2}, \cdots, p_{n,m_n}$$

such that

$$M \subset \sum_{i=1}^{m_n} U_n(p_{n,i}).$$

4. A function f on $M \subset S$ to S' is *uniformly continuous* on M if for each n there is an $m(n)$ such that if $p \subset M$

$$f[MU_{m(n)}(p)] \subset U'_n(f(p)).$$

It is clear that these definitions become the usual ones for metric space with spherical neighborhoods.

The justification for these generalizations is to be sought in the theorems that a set M in a complete space is compact if and only if it is totally bounded, and that if F is uniformly continuous on $M \subset S$ to a complete space S', then there is a function F^* on M to S' identical with F on M and continuous on $\bar{M}$.

Received June 2, 1937; presented to the American Mathematical Society, November, 1935, under the title *Cauchy convergence in non-metric space* (preliminary report), and April, 1937, under the title *Uniform continuity in topological space*.

[1] The relation between completeness and compactness is discussed by J. von Neumann, Annals of Math., vol. 36 (1935). Recent notes on the subject have been published by Garrett Birkhoff, Annals of Math., vol. 38 (1937), pp. 57-60, and by L. M. Graves, Annals of Math., vol. 38 (1937), pp. 61-64.

It will appear that for certain of these results a further restriction must be placed on the spaces, but it does not imply metrizability.

THEOREM 1. *If S is complete and $M \subset S$ is totally bounded, then M is compact.*

Proof. Let p_k be a sequence of points of M. Then

$$(p_k) \subset M \subset \sum_{i=1}^{m_1} U_1(q_{1,i})$$

for some points $q_{1,1}, q_{1,2}, \cdots, q_{1,m_1}$ in S. Denote by $U_1(q_1)$ one of the $U_1(q_{1,i})$ containing a subsequence $p_k^{(1)}$ of p_k. Assume that $U_n(q_n)$ contains a subsequence $p_k^{(n)}$ of p_k. Then

$$(p_k^{(n)}) \subset MU_n(q_n) \subset \sum_{i=1}^{m_{n+1}} U_{n+1}(q_{n+1,i})$$

for some points $q_{n+1,1}, q_{n+1,2}, \cdots, q_{n+1,m_{n+1}}$. Denote by $U_{n+1}(q_{n+1})$ one of the $U_{n+1}(q_{n+1,i})$ which contains a subsequence $p_k^{(n+1)}$ of $p_k^{(n)}$. We now have a sequence of neighborhoods $U_n(q_n)$ and a sequence of sequences $p_k^{(n)}$ such that

$$U_n(q_n) \supset (p_k^{(n)}) \supset (p_k^{(n+1)}).$$

The diagonal sequence $p_k^* = p_k^{(k)}$ satisfies the condition

$$p_k^* \subset U_n(q_n), \qquad k > n.$$

p_k^* is a Cauchy sequence and, S being complete, M is compact.

It is to be noted that the notion of Cauchy sequence is not topologically invariant. This is the case in metric space also. Under uniformly continuous mapping, however, we do have invariance.

THEOREM 2. *If f is uniformly continuous on S to S' and p_k is a Cauchy sequence in S, then $f(p_k)$ is a Cauchy sequence in S'.*

Proof. Consider n and $m(n)$. Then $p_k \subset U_{m(n)}(p)$ for some $p \subset S$ and $k > k_n$. Hence $f(p_k) \subset f[U_{m(n)}(p)] \subset U'_n(f(p))$ for $k > k_n$. This being so for all n, $f(p_k)$ is a Cauchy sequence.

We impose on S the

POSTULATE. *To each $p \subset S$ and positive integer n there is an integer $m(n, p) > 0$ such that if $q \subset U_{m(n,p)}(p)$ then $p \subset U_n(q)$.*

THEOREM 3. *If $M \subset S$ is compact, then M is totally bounded.*

Proof. Assume that M is not totally bounded. Then for some n, M is not contained in the sum of a finite number of $U_n(p)$. For each k and

$$p_1, p_2, \cdots, p_k \subset M$$

there is

$$p_{k+1} \subset M - M \sum_{i=1}^{k} U_n(p_i),$$

M being compact, a subsequence p_{k_t} of p_k has a limit $p \subset S$ so that

$$p_{k_t} \subset U_j(p), \qquad t > t_j.$$

Choosing $j = m(n, p)$ and $t > t_{m(n,p)}$, we have $p \subset U_n(p_{k_t})$ and

$$p_k \subset U_n(p_{k_t})$$

for infinitely many k. Let $p_{k_t} = p_s$ and r be the smallest integer greater than s for which $p_r \subset U_n(p_{k_t})$. We have

$$p_r \subset M - M \sum_{i=1}^{r-1} U_n(p_i) \subset M - M \sum_{i=1}^{s} U_n(p_i) \subset M - MU_n(p_s)$$

and a contradiction.

With S subject to the postulate given above and S' regular and complete, we have

THEOREM 4. *If f is uniformly continuous on $M \subset S$ to S', then there is a function f^* on $\bar{M}$, the closure of M, such that $f = f^*$ on M and f^* is continuous on $\bar{M}$.*

Proof. To each $p \subset \bar{M}$ there is a sequence $p_k \subset M$ such that $\lim p_k = p$. Consider, for p and n, the number $m(n, p)$ of the postulate. Since

$$p_k \subset U_{m(n,p)}(p)$$

for $k_n > k_{m(n,p)}$, $p \subset U_n(p_{k_n})$ for each n and some k_n.

Since f is uniformly continuous on M and $p_k \subset M$, we have

$$f(p_k) \subset f[MU_{m(n)}(p_{k_{m(n)}})] \subset U'_n(f(p_{k_{m(n)}})), \qquad k > k_{m(n)},$$

so that $f(p_k)$ is a Cauchy sequence in S'. S' being complete, there is $p^* \subset S'$ such that $\lim f(p_k) = p^*$. If $\lim q_k = p$ and $q_k \subset M$, there is the sequence $r_k \subset M$ whose terms are alternately p_k and q_k which has p as limit. Hence $\lim f(r_k) = \lim f(q_k) = \lim f(p_k) = p^*$. Thus for each $p \subset \bar{M}$ and every sequence $p_k \subset M$ with limit p, $\lim f(p_k) = p^*$. This defines a single-valued mapping $f^*(p) = p^*$ on $\bar{M}$ to S'.

If $p \subset M$, then $p_k = p \subset M$ and $f^*(p) = \lim f(p_k) = f(p)$ so that $f^* = f$ for $p \subset M$. If f^* is not continuous on $\bar{M}$, there is a $p \subset \bar{M}$ and an n such that for every m, $\bar{M}U_m(p) \supset p_m$ such that $f^*(p_m) \subset S' - U'_n(f^*(p))$. We may choose p_m so that

$$p_m \subset \bar{M}U_{\mu(m)} \subset \bar{M} \prod_{\mu=1}^{m} U_\mu(p).$$

Since S' is regular, there is a ν such that $\bar{U}'_\nu(f^*(p)) \subset U_n(f^*(p))$. Thus for all m

$$f^*(p_m) \subset S' - U'_n(f^*(p)) \subset S' - \bar{U}'_\nu(f^*(p)) \subset S' - U'_\nu(f^*(p)).$$

Now $p_m \subset \bar{M}$ so that there is a sequence $q_{m,k} \subset M$ such that $\lim_k q_{m,k} \subset p_m$. Hence $q_{m,k} \subset \prod_{\mu=1}^{m} U_\mu(p)$ if $k > k_m$. Further $\lim_k f(q_{m,k}) = f^*(p_m)$ implies

$$f(q_{m,k}) \subset S' - \bar{U}'_\nu(f^*(p)) \qquad \text{if } k > k'_m.$$

Hence for $k(m) > \max(k_m, k'_m)$ we have

$$q_{m,k(m)} \subset \prod_{\mu=1}^{m} U_\mu(p), \qquad f(q_{m,k(m)}) \subset S' - \overline{U}'_\nu(f^*(p)) \subset S' - U'_\nu(f^*(p)).$$

Now $\lim_m q_{m,k(m)} = p$, $q_{m,k(m)} \subset M$ and the closure of $S' - U'_\nu(f^*(p))$ yield

$$\lim f(q_{m,k(m)}) = f^*(p) \subset S' - U'_\nu(f^*(p)),$$

which is a contradiction. Thus f^* is continuous on $\overline{M}$.

We give an example of a non-metrizable topological space S satisfying the first denumerability axiom and our postulate. The points of S are the points $p(x, y)$ of the Cartesian plane. Let $S(p, r)$ be the interior of the circle with center p and radius r. The neighborhoods of $p_0(0, 0)$ are defined by

$$U_n(p_0) = S\left(p_0, \frac{1}{n}\right) - \{p\},$$

where $p(x, y)$ has $x > 0$, $y = 0$. If $p(x, y)$ has $y \neq 0$, $U_n(p) = S(p, n^{-1})$. If $p(x, 0)$ has $x > 0$, then $U_n(p) = S[p, (k + n)^{-1}]$, where k is the smallest integer such that the distance from p to p_0 is greater than k^{-1}. The space fails to be metrizable because it is not regular at p_0.

By adding the restriction of regularity to the space we can continue the theory to include Baire's theorems on category and on the distribution of points of continuity of the limit function of a sequence of continuous functions. These theorems are of considerable importance in existence proofs in functional analysis. In the subsequent discussion it will be assumed that the space S is regular as well as complete in the sense defined earlier.

Theorem 5. *If A_n is a sequence of sets in S such that A_{n+1} is dense in A_n and G_n is an open set containing A_n, then $\prod_n G_n$ is dense in A_1.*

Proof. Consider $p_1 \subset A_1$ and any $U_n(p_1)$. There is $\overline{U}_{n_1}(p_1) \subset U_n(p_1)G_1$. Then $U_1(p_1)U_{n_1}(p_1)A_2 \supset p_2$ and there is a neighborhood $U_{n_2}(p_2)$ such that

$$\overline{U}_{n_2}(p_2) \subset U_1(p_1)U_{n_1}(p_1)G_2. \tag{2}$$

If $p_k \subset A_k$ and $U_{n_k}(p_k)$ are defined, the $U_k(p_k)U_{n_k}(p_k)A_{k+1} \supset p_{k+1}$ and there is a $U_{n_{k+1}}(p_{k+1})$ such that

$$\overline{U}_{n_{k+1}}(p_{k+1}) \subset U_k(p_k)U_{n_k}(p_k)G_{k+1}. \tag{$k+1$}$$

From the relations (k), we have

$$p_{k+m} \subset \overline{U}_{n_{k+m}}(p_{k+m}) \subset U_k(p_k), \qquad m \geqq 0.$$

Thus p_k is a Cauchy sequence and, S being complete, $\lim p_k = p$ exists. Also from (k) it follows that

$$p_{k+m} \subset \overline{U}_{n_k}(p_k) \subset G_k, \qquad m \geqq 0,$$

so that $p \subset \overline{U}_{n_k} \subset G_k$ for all k. Thus $p \subset \prod_k G_k U_n(p_1)$ and $\prod_k G_k$ is dense in A.

THEOREM 6. *If A_n is a sequence of G_δ sets in S each dense in every other one then $\prod_n A_n$ is dense in each A_n.*

Proof. Each $A_n = \prod_m G_{m,n}$ where $G_{m,n}$ is open. Let

$$G_1 = G_{1,1}, G_2 = G_{2,1}, \cdots, G_k = G_{m,n}, \cdots$$

be the enumeration of the $G_{m,n}$ by diagonals. If we set $B_k = A_n$ when $G_k = G_{m,n}$, then $G_k \supset B_k$ and B_{k+1} is dense in B_k. Hence $\prod_k G_k = \prod_{m,n} G_{m,n} = \prod_n A_n$ is dense in A_1 by Theorem 5 and consequently in every A_n.

We have a consequence in

THEOREM 7. *If M is a G_δ in S and A_n in a sequence of sets open in M and dense in M, then $\prod A_n$ is dense in M.*

Proof. Each A_n, being open in M, is a G_δ set in S. Each A_n, being dense in M, is dense in every other A_n. Hence by Theorem 6 $\prod_n A_n$ is dense in A_1 and also in M.

From this it follows that S is of the second category. In fact we have

THEOREM 8. *If M is a G_δ in S and A_n is a sequence of subsets of M each nowhere dense in M, then $M - \sum_n A_n$ is dense in M.*

Proof. If $F_n = M\bar{A}_n$ and $G_n = M - F_n$, then G_n is open in M and dense in M. By Theorem 7, $\prod_n G_n = M - \sum_n F_n$ is dense in M, hence $M - \sum_n A_n \supset M - \sum_n F_n$ is dense in M.

This leads to the theorem of Baire on the continuity properties of the limit function of a sequence of continuous functions.

THEOREM 9. *If A is a G_δ in S and $f_n(p)$ is a sequence of functions continuous on A to a metric space R, having the limit function $f(p)$, then the set of points of continuity of $f(p)$ is dense in A.*

Proof. For a given $\eta > 0$, let A_k be the subset of A on which

$$d[f(p), f_n(p)] < \eta \qquad \text{if } n > k.$$

Then $A = \sum_k A_k$ since f is the limit of f_n on A. By Theorem 8, at least one of the A_k is not nowhere dense in A. Hence for some k and some non-empty A^* open in A, $\bar{A}_k \supset A^*$.

Let q be any point of A^*. Then for some $m > k$

$$d[f(q), f_m(q)] < \eta,$$

since $f(q) = \lim f_n(q)$. Now every $U_\mu(q)$ contains a point $p \subset A_k$ and for such p

$$d[f(p), f_n(p)] < \eta, \qquad \text{if } n > k.$$

For the given m and each $n > k$, there is a $U_\nu(q)$ such that

$$d[f_m(p), f_m(q)] < \eta, \qquad d[f_n(p), f_n(q)] < \eta, \qquad \text{if } p \subset U_\nu(q)A,$$

because f_n and f_m are continuous at q. Further, since $m > k$,

$$d[f(p), f_m(p)] < \eta, \qquad \text{if } p \subset A_k.$$

Hence for the given m and each $n > k$ and an existing $p \subset U_\mu(q)U_\nu(q)A_k$, we have

$$d[f(q), f_n(q)] \leqq d[f(q), f_m(q)] + d[f_m(q), f_m(p)] + d[f_m(p), f(p)] + d[f(p), f_n(p)] + d[f_n(p), f_n(q)] < 5\eta.$$

Now consider any $\epsilon > 0$, $\eta < \frac{1}{5}\epsilon$ and the corresponding A^*. Since A^* is open in A, there is, for each $q \subset A^*$, a neighborhood $U_\epsilon(q)$ such that

$$d[f(p), f_n(p)] < \epsilon, \qquad \text{if } n > k,\ p \subset U_\epsilon(q)A.$$

Let $A(\epsilon)$ be the set of all q in A satisfying this condition. Then $A(\epsilon)$ is open in A and not empty. The function $f(p)$ is continuous at each point of $\prod_n A(n^{-1})$. These statements remain true if the set A is replaced by any set B open in A. Hence $A(n^{-1})$ is dense in A for each n and, by Theorem 7, $\prod_n A(n^{-1})$ is dense in A. Thus the set of points of A at which $f(p)$ is continuous is dense in A.[2]

As an example of a regular Hausdorff space which satisfies the first denumerability axiom and is complete but not metrizable we have the space T whose points are the ordinals of the first and second classes and whose neighborhoods are open segments of the ordered set T.

UNIVERSITY OF KENTUCKY.

[2] It is to be noted that the proofs of Theorems 5-9 are parallel to those for metric spaces. Cf. Hausdorff, *Grundzüge der Mengenlehre*.

SOLUTIONS OF SYSTEMS OF DIFFERENTIAL EQUATIONS IN TERMS OF INFINITE SERIES OF DEFINITE INTEGRALS

By Jesse Pierce

Introduction. The general solution of a differential equation of the first order and first degree can be found in terms of an infinite series of definite integrals.[1] The definite integrals appearing in the solution are solutions of linear differential equations.

The same method is applicable to finding the general solution of a system of differential equations. The functions giving the solution, however, are expressed in terms of infinite series of solutions of systems of linear differential equations. The general solution of each system of linear differential equations can be found in terms of infinite series of definite integrals.[2]

Hence the solution of the original system of differential equations is expressed in terms of infinite series of infinite series of definite integrals except in the case where the original system is linear or the system comprises just one differential equation of the first order.

In the present paper the general solution of a system of differential equations will be found in terms of infinite series of definite integrals in which every integrand consists of a finite number of terms.

The system of differential equations to be considered has the form

$$\frac{dy_i}{dt} = \varphi_i(t, y_1, \cdots, y_n) \qquad (i = 1, \cdots, n), \tag{1}$$

where the $\varphi_i(t, y_1, \cdots, y_n)$ are analytic in the y_j $(j = 1, \cdots, n)$ and have as coefficients $a_{i\mu_1\cdots\mu_n}$ functions of t which are integrable (Riemann) and satisfy the inequalities

$$| a_{i\mu_1\cdots\mu_n} | \leqq f(u), \tag{2}$$

$f(u)$ being a positive integrable function of u, the arc length of a rectifiable curve drawn from the origin to the point t. The exponent of y_j $(j = 1, \cdots, n)$ in the expanded form of the φ_i is represented by μ_j. The independent variable t is assumed to be of the form

$$t = \varphi(u) + \sqrt{-1}\,\psi(u), \tag{3}$$

where $\varphi(u)$, $\psi(u)$ are real functions with continuous first derivatives.

Received June 9, 1937.

[1] *Solutions of a differential equation of the first order and first degree in terms of infinite series of definite integrals*, presented by the author to the Ohio Section of the Mathematical Association of America, April, 1937.

[2] *Solutions of systems of linear differential equations in the vicinity of singular points*, American Mathematical Monthly, vol. 43 (1936), pp. 530–539.

Each unknown y_i is obtained as the sum of a series y_{ih} $(h = 1, 2, \cdots)$, whose terms are found by integrating, sequentially, certain polynomials in the y_{jk} $(j = 1, \cdots, n; k = 1, \cdots, h - 1)$.

For convenience, the sequence $\mu_1, \cdots, \mu_n$ will be represented by μ and hence the coefficient $a_{i\mu_1 \cdots \mu_n}$ will be represented by $a_{i\mu}$.

In §1 the formal solution of the system of differential equations (1) is found and the convergence of this solution is proved in §2. In §3 a more general system of differential equations is reduced to the form (1) by a simple transformation.

1. **Formal solution of the system of differential equations** (1). Equations (1) can be written in the expanded form

$$\frac{dy_i}{dt} = a_i(t) + \sum_{j=1}^{n} a_{ij}(t)\cdot y_j + \sum_{\nu=2}^{\infty} a_{i\mu}(t) y_1^{\mu_1} y_2^{\mu_2} \cdots y_n^{\mu_n} \qquad (i = 1, \cdots, n), \tag{4}$$

where $\nu = \mu_1 + \cdots + \mu_n$.

When the y_i, in equations (4), are replaced by

$$y_i = \sum_{h=1}^{\infty} y_{ih} K^h, \tag{5}$$

there results

$$\sum_{h=1}^{\infty} K^h \frac{dy_{ih}}{dt} = a_i(t) + \sum_{j=1}^{n} a_{ij}(t) \sum_{h=1}^{\infty} K^h y_{jh} + \sum_{h=2}^{\infty} K^h \mathcal{F}_{ih}(t, y_{jk}) \tag{6}$$
$$(i = 1, \cdots, n; k = 1, \cdots, h - 1),$$

where the $\mathcal{F}_{ih}(t, y_{jk})$ are polynomials in the y_{jk} of degree h and each coefficient is one of the $a_{i\mu}(t)$. The parameter K is introduced in order that a convenient arrangement of the terms in the right-hand members of equations (6) can be made.

A formal solution of equations (6) can be found by replacing K by unity and then solving the following system of differential equations:

$$\begin{cases} \dfrac{dy_{i1}}{dt} = a_i(t), \\[2ex] \dfrac{dy_{i2}}{dt} = \displaystyle\sum_{j=1}^{n} a_{ij}(t) y_{j1}, \\[2ex] \dfrac{dy_{ih}}{dt} = \displaystyle\sum_{j=1}^{n} a_{ij}(t)\cdot y_{j,h-1} + \mathcal{F}_{i,h-1}(t, y_{jk}) \qquad (h = 2, 3, \cdots). \end{cases} \tag{7}$$

Equations (7) are obtained by equating the coefficient of K^h in the left-hand member of (6) to the coefficient of K^{h-1} in the right-hand member.

Equations (7) have the formal solution

$$(8)\qquad \begin{cases} y_{i1} = \displaystyle\int_0^t a_i(t)\,dt + c_i \equiv \eta_{i1}(t), \\ y_{i2} = \displaystyle\int_0^t \sum_{j=1}^n a_{ij}(t)\,\eta_{j1}(t)\,dt \equiv \eta_{i2}(t), \\ y_{ih} = \displaystyle\int_0^t \left[\sum_{j=1}^n \theta_{ij}(t)\,\eta_{j,h-1}(t) + \mathcal{F}_{i,h-1}(t, \eta_{jk}(t))\right] dt \equiv \eta_{ih}(t) \\ \qquad\qquad (k = 1, \cdots, h-2;\ h = 2, 3, \cdots), \end{cases}$$

where the c_i are arbitrary parameters.

The set of functions defined by the infinite series

$$(9)\qquad y_i = \sum_{h=1}^{\infty} \eta_{ih}(t) \qquad (i = 1, \cdots, n),$$

is a formal solution of the system of differential equations (4).

2. **Proof of the convergence of the series** (9). A particular case of the system of differential equations (1) is the system

$$(10)\qquad \frac{dY_i}{du} = \frac{f(u)}{1 - \sum_{j=1}^n Y_j} \qquad (i = 1, \cdots, n),$$

whose right-hand members dominate those of (1). The system (10) has a solution in which all of the Y_i are equal, that is, the Y_i all satisfy the equation

$$(11)\qquad \frac{dY}{du} = \frac{f(u)}{1 - nY}.$$

The general solution of the differential equation (11) is

$$(12)\qquad Y = \frac{1 - (1 - 2n[G(u) + c])^{\frac{1}{2}}}{n},$$

where

$$(13)\qquad G(u) = \int_0^u f(u)\,du,$$

and c is an arbitrary parameter. We shall consider c real and non-negative. The minus sign is used before the radical in order that all of the terms in the right-hand member of (12), when expanded in a power series in $[G(u) + c]$, be positive. This expansion has the form

$$(14)\quad Y = [G(u) + c] + \frac{n[G(u) + c]^2}{2} + \frac{n^2[G(u) + c]^3}{2} + \cdots = b_h n^{h-1}[G(u) + c]^h,$$

where the b_h are real positive constants. The series (14) will converge when

$$2n[G(u) + c] < 1. \tag{15}$$

The power series (14) can be found directly from the differential equation (11) by expanding the right-hand member of (11) in a power series in Y and then making the substitution

$$Y = \sum_{h=1}^{\infty} Y_h K^h. \tag{16}$$

The result can be arranged in the form

$$\begin{aligned}\sum_{h=1}^{\infty} \frac{dY_h}{dt} K^h &= f(u) + f(u) \sum_{h=1}^{\infty} nY_h K^h \\ &+ f(u) \sum_{h=2}^{\infty} K^h [n^2\{Y_1 Y_{h-1} + \cdots + Y_{h-1}Y_1\} + \cdots + n^h Y_1^h] \\ &= f(u) + f(u) \sum_{h=1}^{\infty} K^h nY_h + \sum_{h=2}^{\infty} K^h F_h(u, Y_k) \quad (k = 1, \cdots, h-1).\end{aligned} \tag{17}$$

It follows from the inequality (2) that the polynomial $F_h(u, Y_k)$ dominates the polynomial $\mathcal{F}_{ih}(t, y_{jk})$.

The system of differential equations corresponding to (7) is

$$\begin{cases} \dfrac{dY_1}{du} = f(u), \\ \dfrac{dY_2}{du} = f(u)nY_1, \\ \dfrac{dY_3}{du} = f(u)[nY_2 + n^2 Y_1^2], \\ \dfrac{dY_h}{du} = f(u)[nY_{h-1} + F_{h-1}(u, Y_k)] \quad (k = 1, \cdots, h-2;\ h = 3, 4, \cdots). \end{cases} \tag{18}$$

Equations (18) can be solved in terms of indefinite integrals in the form

$$\begin{cases} Y_1 = \displaystyle\int f(u)\, du = G(u) + c \equiv H_1(u), \\ Y_2 = \displaystyle\int f(u)nH_1(u)\, du = \tfrac{1}{2}n[G(u) + c]^2 \equiv H_2(u), \\ Y_3 = \displaystyle\int f(u)\{nH_2(u) + n^2H_1^2(u)\}\, du = \tfrac{1}{2}n^2[G(u) + c]^3 \equiv H_3(u), \\ Y_h = \displaystyle\int [f(u)nH_{h-1}(u) + F_{h-1}(u, H_k(u))]\, du = \alpha_h[G(u) + c]^h \equiv H_h(u), \end{cases} \tag{19}$$

where the α_h are positive constants.

The series

$$Y = \sum_{h=1}^{\infty} H_h(u) \tag{20}$$

is a power series in $[G(u) + c]$ and hence is precisely the same as the series **(14)**.

When equations (18) are solved as in §1 in terms of definite integrals, one obtains

$$\tag{21} \begin{cases} Y_1 = \displaystyle\int_0^u f(u)\,du + c = G(u) + c \equiv Z_1(u), \\ Y_2 = \displaystyle\int_0^u f(u)nZ_1(u)\,du = \tfrac{1}{2}n[G(u) + c]^2 - \tfrac{1}{2}nc^2 \equiv Z_2(u), \\ Y_3 = \displaystyle\int_0^u f(u)[nZ_2(u) + n^2Z_1^2(u)]\,du = \tfrac{1}{2}n^2[G(u) + c]^3 - \tfrac{1}{2}n^2c^2G(u) - \tfrac{1}{2}n^2c^3 \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \equiv Z_3(u), \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots . \end{cases}$$

By comparing equations (21) with (19) it is clear that

$$Z_1(u) \equiv H_1(u), \qquad Z_h(u) \leqq H_h(u) \qquad (h = 2, 3, \cdots). \tag{22}$$

It follows from equations (21) and (8) that

$$|\eta_{ih}(t)| \leqq Z_h(u) \qquad (h = 1, 2, \cdots), \tag{23}$$

provided

$$|c_i| \leqq c. \tag{24}$$

Hence the series (9) will converge when the inequalities (15) and (24) are satisfied.

Since the series (9) are absolutely convergent and the independent variable t satisfies the relation (3), the functions y_i have derivatives of the first order and satisfy the system of differential equations (4) for all values of t on the path of integration except at the set of points of measure zero at which the coefficients $\alpha_{i\mu}$ are discontinuous.

3. **Systems of differential equations with more general coefficients than those of** (4). Consider the system of differential equations

$$\frac{dx_i}{dt} = \theta_i(t) + \theta(t)x_i + \sum_{j=1}^{n} \theta_{ij}(t)x_j + \sum_{\nu=2}^{\infty} f_{i\mu}(t)x_1^{\mu_1}x_2^{\mu_2}\cdots x_n^{\mu_n} \quad (i = 1, \cdots, n), \tag{25}$$

where μ represents the sequence $\mu_1, \cdots, \mu_n$ and $\nu = \mu_1 + \cdots + \mu_n$. The function $\theta(t)$ is assumed to have the indefinite integral

$$\int \theta(t)\,dt = \beta(t), \tag{26}$$

which may become infinite at a set of points of measure zero on the path of integration defined in §1. The functions $\theta_i(t)$ have the form

$$\theta_i(t) = \zeta_i(t)e^{\beta(t)}, \tag{27}$$

where the $\zeta_i(t)$ are integrable (Riemann) and satisfy the inequalities

$$|\zeta_i(t)| \leqq f(u), \tag{28}$$

for all values of t on the path of integration. The functions $\theta_{ij}(t)$ are integrable and satisfy the inequalities

$$|\theta_{ij}(t)| \leqq f(u), \tag{29}$$

for all values of t on the path of integration. The functions $f_{i\mu}(t)$ have the form

$$f_{i\mu}(t) = \zeta_{i\mu}(t)e^{-(\nu-1)\beta(t)} \qquad (\nu = 2, 3, \cdots), \tag{30}$$

where the $\zeta_{i\mu}(t)$ are integrable, and satisfy the inequalities

$$|\zeta_{i\mu}(t)| \leqq f(u), \tag{31}$$

for all values of t on the path of integration.

The transformation

$$x_i = e^{\beta(t)}y_i, \tag{32}$$

reduces the system of differential equations (25) to the form (4) and hence the system (25) has the solution

$$x_i = e^{\beta(t)} \sum_{h=1}^{\infty} \eta_{ih}(t), \tag{33}$$

where the $\eta_{ih}(t)$ are defined by equations (8).

When the real part of the function $\beta(t)$ approaches minus infinity as t approaches t', a point on the path of integration, then

$$\lim_{t \to t'} x_i(t) = 0. \tag{34}$$

When the real part of the function $\beta(t)$ approaches plus infinity as t approaches t', then

$$\lim_{t \to t'} x_i(t) = \infty, \tag{35}$$

provided

$$\lim_{t \to t'} \sum_{h=1}^{\infty} \eta_{ih}(t) \neq 0. \tag{36}$$

Conclusion. The coefficients of the differential equations (4) and (25) can satisfy the assumptions of §§1 and 3 without being continuous or bounded with respect to t.

The solutions (9) and (33) are the general solutions of the systems (4) and (25), respectively, because they contain n parameters c_i, which are arbitrary except for the inequalities (15) and (24).

In the case where the limit of the real part of the function $\beta(t)$ is minus infinity when t approaches zero, the initial values of the dependent variables x_i are zero for every set of values of the c_i.

HEIDELBERG COLLEGE.

NON-n-ALTERNATING TRANSFORMATIONS

By D. W. Hall and G. E. Schweigert

Let A and B be compact metric spaces and $T(A) = B$ a single-valued continuous transformation. We shall say that T is *non-n-alternating* provided that, for any point x of B for which there exists a cutting K of $A - T^{-1}(x)$ consisting of at most n points, there is no point y of B such that $T^{-1}(y)$ intersects both sets of the separation $A - (T^{-1}(x) + K) = A_1 + A_2$. If K is the null set, this is the definition of a non-alternating transformation.[1] Consequently, this type of transformation is non-alternating; in fact, we have the following characterization:

Theorem I. *A necessary and sufficient condition that a single-valued continuous transformation $T(A) = B$ be non-n-alternating is that T be non-alternating on the complement of every subset of A consisting of at most n points.*

Proof. Let x and y be points of B and K any subset of A consisting of at most n points. If $T^{-1}(x)\cdot(A - K)$ separates[2] $T^{-1}(y)\cdot(A - K)$ in $A - K$, i.e., if $(A - K) - T^{-1}(x)\cdot(A - K) = A_1 + A_2$, $T^{-1}(y)\cdot(A - K)\cdot A_i \neq 0$ $(i = 1, 2)$, then this separation may be written in the form $(A - T^{-1}(x)) - K = A_1 + A_2$. Hence K separates $T^{-1}(y)$ in $A - T^{-1}(x)$, contrary to the definition of non-n-alternating. Thus the condition is necessary.

To establish the sufficiency, we notice that if there exist two points x, y in B and a cutting K of $A - T^{-1}(x)$ consisting of at most n points such that $T^{-1}(y)$ intersects both the sets A_1 and A_2 of the separation $A - (T^{-1}(x) + K) = A_1 + A_2$, then $(A - K) - T^{-1}(x)\cdot(A - K) = A_1 + A_2$ and therefore $T^{-1}(y)\cdot(A - K)$ is separated by $T^{-1}(x)\cdot(A - K)$ in $A - K$. Consequently, T is not non-alternating on $A - K$. This proves the sufficiency.

Lemma. *If $T(A) = B$ is non-n-alternating, B is non-degenerate, $y \in B$, and two points of $T^{-1}(y)$ are separated in A by a cutting K consisting of $k \leqq n + 1$ points, then $k = n + 1$ and $T(K) = y$.*

Proof. If $k \leqq n$, then T is non-alternating on the complement of K, by Theorem I. But this is impossible since $T^{-1}(y)$ intersects two components of this complementary set. Thus $k = n + 1$. If $T(K) \neq y$, there exists a point p in K such that $T(p) \neq y$. Then the set of n points $(K - p)$ separates $T^{-1}(y)$ in $A - T^{-1}(T(p))$, contrary to the fact that T is non-n-alternating. Therefore, $T(K) = y$.

One consequence of this lemma, namely, the fact that a point of order not

Received June 12, 1937.

[1] See G. T. Whyburn, *Non-alternating transformations*, American Journal of Mathematics, vol. 56 (1934), pp. 294–302.

[2] If L and M are subsets of N, we say that L "separates" M in N provided M is contained in $N - L$ and $N - L = N_1 + N_2$, where $N_1\bar{N}_2 = 0 = \bar{N}_1N_2$ and $MN_1 \neq 0 \neq MN_2$.

greater than n in A is equal to the inverse of its image provided B is non-degenerate, suggests the following theorem. This theorem covers, as a special case, the class of regular curves of order $n + 1$.

THEOREM II. *If A is a compact metric space every pair of points of which is separated by a set containing at most $(n + 1)$ points and $T(A) = B$ is non-n-alternating $(n > 1)$, then T is a homeomorphism on A provided B is non-degenerate.*

Proof. We first show that the inverse of every point b of B is an A-set.[3] Let $G = T^{-1}(b)$; then G is closed and lies in a single component H of A. From the hypotheses of the theorem it follows that H is a locally connected continuum. Consequently, if p and q are any two points of G, there is a simple arc pq joining these two points in H. If we assume that there is a point x in pq which is not in G, it follows that there is a last point y of G in the arc from p to x and a first point z of G in the arc from x to q. Thus the open arc yxz contains no points of G. By hypothesis, there exists a cutting K consisting of at most $(n + 1)$ points and separating y and z in A, so that some point of K must lie on the open arc yxz. Hence $T(K)$ is not b, contrary to the lemma. It follows that G contains every simple arc joining two of its points, and hence, being closed, it is an A-set.

It is also a consequence of the lemma that no set of n points separates G, so that this set lies in a true cyclic element C of A; therefore $G = C$, since C is a minimal A-set.[4] The set C now has the property that any pair of its points can be irreducibly separated by $(n + 1)$ points. It follows[5] that C cannot exist except as a single point. Thus T is 1-1 and hence a homeomorphism. This completes the proof of the theorem.

Under the same hypotheses for $n = 1$ the above proof holds except for the non-existence of the true cyclic element C. If C exists, it must be a simple closed curve, and the image space B is not only a boundary curve,[6] but it is homeomorphic with the original curve, provided that no true cyclic element of A has a degenerate image.

If $n = 0$, the transformation is non-alternating and acts on a space having dendrites as components; this situation has already been discussed in the original paper on non-alternating transformations.[1]

[3] That is, $T^{-1}(b)$ is closed and contains every simple arc joining two of its points in A. See Kuratowski and Whyburn, Fundamenta Mathematicae, vol. 16, p. 309.

[4] See Kuratowski and Whyburn, loc. cit.

[5] This follows from the theorem that for $n > 2$ there exists no continuous curve M such that for every pair of points P and Q of M there are exactly n independent arcs from P to Q, but there does not exist for any pair of points $(n + 1)$ such arcs. This theorem has been proved for the cases $n = 3$ and $n = 4$ by Kusner and published in the Comptes Rendus des Séances de la Société des Sciences de Varsovie (1932). It has been established for the general case, but not yet published, by J. R. Kline.

[6] That is, a compact locally connected continuum each true cyclic element of which is a simple closed curve. This term has been suggested by G. T. Whyburn, loc. cit., p. 301. It follows that a boundary curve is a compact continuum every pair of points of which is separated by at most two points. Here, this characterization is needed rather than the definition.

From this same source we learn that the property of being a boundary curve is invariant under non-alternating transformations. Since this result represents in a certain sense a sharper form of the present theorem, it might be suspected that for a compact metric space A the property of being separated between each pair of its points by at most $(n + 2)$ points is invariant under

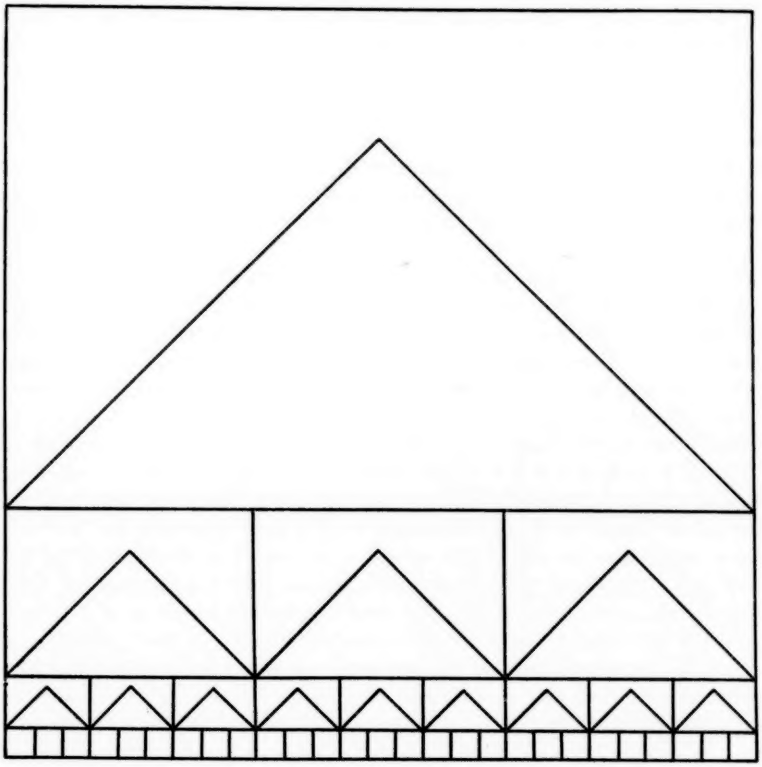

FIG. 1

non-n-alternating transformations. That such is not the case can be shown by simple examples. However, the failure of the purest form of the analogy by no means denies the possibility of other closely connected results. For example: Is the non-n-alternating image of a regular curve of order $(n + 2)$ hereditarily locally connected?

The question as to whether or not the property of being a regular curve is

invariant under non-alternating transformations was raised by G. T. Whyburn and offered difficulties which led to this and to at least one other specialized type of non-alternating transformation. In answer to this question we may now make the following statement: *There exists a regular curve A of order three and a non-alternating transformation*[7] $T(A) = B$ *such that the inverse of any point of B is at most three points, but B is not hereditarily locally connected, hence certainly not a regular curve.*

Example. Let P_0 denote a square on (0, 1) and for each positive integer k let P_k be the square on (0, 1) which lies above and makes an angle $\theta_k = \frac{\pi}{2^k}$ with P_0. Consider, for future reference, the intervals with end points $3^{-m}r + 3^{-m-1}$ and $3^{-m}r + 2 \cdot 3^{-m-1}$ $(m = 0, 1, \cdots ; r = 0, 1, \cdots , 3^m - 1)$, and denote each such interval by $I_{m,r}$. In P_k for $k = 3^m + r$, make the following construction: (i) on the base $I_{m,r}$ erect a square and denote the side opposite the base by s; (ii) using s as the middle third of the hypotenuse construct an isosceles right triangle which is disjoint with the interior of this new square.

The space A is the sum of the boundary of P_0 and all the figures constructed in (i), (ii) above. The transformation $T(A) = B$ will consist of a simple identification of all the squares P_k with the particular square P_0, i.e., let $\theta_k = 0$ for all k. The image space B may be constructed as indicated by the figure. We have at once that T is one-to-one on A except at the ends of the hypotenuse of each triangle constructed in (ii) and at certain obvious points on the lines perpendicular to the basic unit interval. If p is any point of A such that $T^{-1}(T(p)) \neq p$, then p is not a point of the basic unit interval. Consequently, p lies in a non-degenerate cyclic element of $A \cdot P_k$, where $p \in P_k$. No other point of $T^{-1}(T(p))$ is in P_k, hence $T^{-1}(T(P))$ does not separate $A \cdot P_k$. It follows easily that no inverse set separates A, i.e., T is non-separating, hence surely non-alternating. It will also be observed that the bases of all the triangles constructed on squares of the same height are joined in B to form an interval of unit length, and that the sequence of intervals thus formed has the unit interval as a continuum of convergence. Thus B is not hereditarily locally connected. The remaining properties of T, A, and B are easy consequences of their respective definitions.

UNIVERSITY OF VIRGINIA.

[7] In fact, this transformation is *non-separating* in the sense of Wardwell, i.e., for no $b \in B$ does $T^{-1}(b)$ separate A. See James F. Wardwell, *Non-separating transformations*, this Journal, vol. 2 (1936), pp. 745-750.

RESIDUATION IN STRUCTURES OVER WHICH A MULTIPLICATION IS DEFINED

By Morgan Ward

I. Introduction

1. Consider a set of elements A, B, C, $\cdots$ forming an abstract structure[1] Σ over which there is defined a commutative and associative multiplication XY. The multiplication operation is connected with the structure by assuming that it is distributive with respect to union:

$$A(B, C) = (AB, AC) \qquad \text{for } A, B, C \text{ in } \Sigma. \tag{3.3}$$

This condition is satisfied in the important instance of the ideals of a commutative ring.

If we assume that any subclass of elements of Σ has a union and correspondingly strengthen assumption (3.3), the existence of a *residual* (German, *idealquotient*[2]) $A:B$ follows for each pair of elements A, B of Σ with the defining properties

$$A \supset (A:B)B; \qquad \text{if } A \supset XB, \text{ then } A:B \supset X.$$

It is easy to show that the residual thus defined has the formal properties of the residual in polynomial ideal theory (Macaulay, [3]). But in the special instance of ordinary arithmetic (when Σ is interpreted as the ring of rational integers) the residual has a number of additional interesting properties which do not hold in general; for example,

$$(A:B, B:A) = I,^3 \qquad (A, B):M = (A:M, B:M),$$

$$M:[A, B] = (M:A, M:B), \qquad M:AB = \{(M:A)(M:B)\}:M,$$

$$A:(B:A) = A(A, B):B, \qquad [A, B]:(A, B) = (A:B)(B:A).$$

The problem arises then of determining the conditions under which these and other properties of residuation in ordinary arithmetic will hold in the abstract structure. It is not difficult to show that it suffices to assume[4]

Postulate E. *If A divides B, there exists a unique element Q such that $AQ = B$.*

Received June 22, 1937.

[1] Other terms are "dual group", "Verband", "lattice". For a definition, see §2 of this paper, or O. Ore, reference [1] at the close of the paper.

[2] The concept appears to be due to Dedekind [4]. See van der Waerden [2] or Macaulay [3].

[3] I here is the unit element with respect to multiplication. See §§2 and 3.

[4] Postulate E is satisfied in every principal ideal ring.

But this solution of the problem is trivial. First of all, Postulate E is far stronger than necessary. Secondly, even if we weaken E by no longer requiring Q to be unique, the assumption is still undesirable because it renders the concept of the residual superfluous; $A:B$ is merely a particular one of the quotients[5] $\dfrac{A}{(A, B)}$. We assume accordingly the weaker

POSTULATE C. *If A divides B, there exists an element P such that* $A = B:P$.

We shall show here that Postulate C alone suffices to prove that $(A:B, B:A) = I$, $(A, B):M = (A:M, B:M)$ and $M:[A, B] = (M:A, M:B)$. Postulate C is a sufficient, but not necessary, condition that Σ be an arithmetic structure, and a necessary condition that Σ be a Boolean algebra, if multiplication is identified with cross-cut.

If we interpret our multiplication as the cross-cut operation of the structure (so that the structure is arithmetic by assumption 3.3), we obtain a residuation operation within the structure which does not seem to have been investigated even in common arithmetic.

It is also possible to define residuation abstractly over the structure by a proper selection of the properties of residuation given in §4 without any reference to multiplication. This investigation has been carried out by R. P. Dilworth in an unpublished paper.

A complete postulational analysis of the interrelationships between the structure properties, and the operations of residuation and multiplication would appear desirable, but will not be given here. We shall content ourselves with an informal treatment, introducing our postulates on the basis of naturalness and convenience. Their consistency will be evident.

2. We begin by recalling briefly the defining properties of a structure (Ore, [1]). We postulate the existence of a well defined division relation $\supset$ which is transitive and reflexive. The equality relation $=$ is then defined in terms of $\supset$ by $A = B$ if and only if $A \supset B$ and $B \supset A$. For any two elements A and B of Σ we postulate the existence of elements D and M such that

$$D \supset A, D \supset B; \quad \text{if} \quad X \supset A, X \supset B, \quad \text{then} \quad X \supset D.$$

$$A \supset M, B \supset M; \quad \text{if} \quad A \supset Y, B \supset Y, \quad \text{then} \quad M \supset Y.$$

D and M are called the union and cross-cut of A and B. They are determined up to equal elements. We write[6] $D = (A, B)$ and $M = [A, B]$.

If an element E divides every element T of a fixed subclass Θ of Σ, we write

[5] We define a quotient $Q = \dfrac{A}{B}$ as an element Q such that $A = QB$. See §4.

[6] Ore uses $[A, B]$ for union and (A, B) for cross-cut. We prefer to retain as far as possible the notation and terminology of ideal theory.

$E \supset \Theta$. We state explicitly our assumption of closure of Σ with respect to union (Ore [1], p. 409):

Postulate A. *For every subclass Θ of elements of Σ, there exists an element U called the union of Θ such that*

$$U \supset \Theta; \quad \text{if} \quad X \supset \Theta, \quad \text{then} \quad X \supset U.$$

We write $U = u(\Theta)$.

In particular the element $u(\Sigma)$ divides every element of Σ. We shall call it the *identity* element of Σ and denote it by I.

If we assume the existence of a null element divisible by every other element, then the closure of Σ with respect to union obviously entails the closure of Σ with respect to cross-cut. Conversely, closure with respect to cross-cut and the existence of an identity imply closure with respect to union.

If for every three elements A, B, C of Σ

$$(A, [B, C]) = [(A, B), (A, C)], \tag{2.1}$$

the structure is said to be arithmetic, or distributive. (2.1) is equivalent to $[A, (B, C)] = ([A, B], [A, C])$.

We recall that $A \supset B$ if and only if $A = (A, B)$ and $B = [A, B]$, and that union and cross-cut are associative, commutative, and idempotent operations.

3. We next assume that Σ is closed with respect to an associative and commutative operation $X \cdot Y$ or XY:

(3.1) $A \cdot B$ is in Σ if A, B are in Σ; $A \cdot (B \cdot C) = (A \cdot B) \cdot C$; $A \cdot B = B \cdot A$.

We call this operation multiplication. We make the further assumptions

(3.2) $I \cdot A = A$ for every element A of Σ.

Here I is the identity element of the structure.

(3.3) $A \cdot (B, C) = (A \cdot B, A \cdot C)$ for any three elements A, B, C of the structure.

The following rules are easy consequences of these assumptions:

$B \supset C$ implies $AB \supset AC$. $A \supset B$ and $C \supset D$ imply $AC \supset BD$.

$A = BC$ implies $B \supset A$. $[A, B] \supset AB \supset [A, B](A, B)$.

$(A, B) = I$ implies $AB = [A, B]$ and $(A, BC) = (A, C)$, any C.

On account of Postulate A, assumption (3.3) must be strengthened, as from (3.3) we can merely deduce the distributivity of multiplication with respect to union for a *finite* number of elements. If Θ and Φ are subclasses of Σ, we define their product $\Theta\Phi$ as the subclass of all products of elements of Θ and elements of Φ If Θ consists of a single element T, we write $T\Phi$ for $\Theta\Phi$. We assume[7]

[7] It suffices to assume merely that $u(T\Phi) = Tu(\Phi)$ for the developments which follow; but this apparently weaker assumption is easily shown to be equivalent to Postulate B.

POSTULATE B. *The union of the product of any two subclasses of Σ is the product of their unions.*

II. Residuation

4. The element $R = A:B$ is called *the residual* of B with respect to A if

$$(4.1) \qquad A \supset RB; \qquad A \supset XB \text{ implies } R \supset X.$$

The residual always exists for any A, B of Σ. For since $A \supset AB$, the class Θ of all elements X such that $A \supset XB$ is non-empty. Let $R = u(\Theta)$. Then by Postulate B, since $A \supset \Theta B$, $A = u(A) \supset u(\Theta B) = u(\Theta)u(B) = RB$, or $A \supset RB$. And if $A \supset XB$, X lies in Θ, so that $R \supset X$.

The following properties of residuation which are needed later follow exactly as in the ideal theory of commutative rings (van der Waerden [2], Chapter XII; Macaulay [3], Chapter 3):

$$(4.2) \quad A:A = I, \qquad A:I = A, \qquad I \supset A:B \supset A.$$

$$(4.21) \quad A:B = A:(A, B) = [A, B]:B.$$

$$(4.3) \quad A:B = I, \text{ if and only if } A \supset B.$$

$$(4.31) \quad \text{If } A:B = A, \text{ then } A \supset XB \text{ implies } A \supset X.$$

$$(4.4) \quad (A:B):C = (A:C):B = A:BC.$$

$$(4.41) \quad M:A = B \text{ implies } AB:A = B.$$

$$(4.5) \quad M:(M:N) \supset N.$$

$$(4.51) \quad A \supset B \text{ implies } M:B \supset M:A \text{ and } A:M \supset B:M.$$

THEOREM 4.1 (Macaulay [3], p. 32). *If M and N are any two elements of Σ and $A = M:N$, $B = M:(M:N)$, then $B = M:A$, $A = M:B$.*

A and B are then said to be *mutually residual* with respect to M.

The *quotient* $Q = \dfrac{A}{B}$ of two elements A and B of Σ is defined by

$$(4.6) \qquad A = QB; \quad \text{if} \quad A = XB, \quad \text{then} \quad Q \supset X.$$

Unlike the residual, the quotient need not exist even if $B \supset A$. But if the class Θ of all X such that $A = XB$ is non-empty, $\dfrac{A}{B}$ exists and is the union $u(\Theta)$. Clearly $\dfrac{A}{A} = I$, $\dfrac{A}{I} = A$, $\dfrac{AB}{A}$ exists.

THEOREM 4.2. *If the quotient $\dfrac{A}{B}$ exists, it equals the residual $A:B$.*

Proof.[8] Let $\dfrac{A}{B} = V$, $A:B = W$. Then by (4.1), (4.6) $A = BV \rightarrow$

[8] We use when convenient $\rightarrow$ and $\sim$ for formal implication and equivalence to shorten the proofs.

$A \supset BV \rightarrow W \supset V \rightarrow BW \supset BV \rightarrow BW \supset A \rightarrow BW = A \rightarrow V \supset W$. $W \supset V$ and $V \supset W \rightarrow W = V$.

It follows from formula (4.21) that $A:B = \dfrac{A}{(A, B)} = \dfrac{[A, B]}{B}$ whenever the indicated quotients exist. Consequently, if we assume

Postulate D. *If A divides B, there exists at least one element Q such that*

$$A = QB,$$

then the residual $A:B$ reduces to the quotient $\dfrac{A}{(A, B)}$. We shall not assume Postulate D in this paper.

We conclude this division of the paper with a lemma which we shall need subsequently.

Lemma 4.1. *If $(R, S) = I$, then $(P, Q) \supset [P:R, Q:S]$.*

Proof. Let $P:R = U$, $Q:S = V$. By (4.1), $P \supset UR$, $Q \supset VS$ so that $(P, Q) \supset (UR, VS)$. It suffices then to show that $(UR, VS) \supset [U, V]$. $(R, S) = I \rightarrow [U, V] = [U, V](R, S) = (R[U, V], S[U, V])$. Now $U \supset [U, V] \rightarrow UR \supset R[U, V]$; $V \supset [U, V] \rightarrow VS \supset S[U, V]$. Hence $(UR, VS) \supset (R[U, V], S[U, V]) \supset [U, V]$.

III. Distributive properties of residuation

5. The following two "distributive laws" for residuation (Macaulay [3], p. 33) are due in essence to Dedekind [4]:

I $\qquad M:(A_1, A_2, \cdots, A_n) = [M:A_1, M:A_2, \cdots, M:A_n]$,

II $\qquad [A_1, A_2, \cdots, A_n]:M = [A_1:M, A_2:M, \cdots, A_n:M]$.

In common arithmetic the residual satisfies the additional distributive laws

III $\qquad (A_1, A_2, \cdots, A_n):M = (A_1:M, A_2:M, \cdots, A_n:M)$,

IV $\qquad M:[A_1, A_2, \cdots, A_n] = (M:A_1, M:A_2, \cdots, M:A_n)$.

Since III and IV are easily proved by induction from the case $n = 2$, we shall discuss them here in the form

$$(A, B):M = (A:M, B:M), \tag{5.1}$$

$$M:[A, B] = (M:A, M:B). \tag{5.2}$$

That III and IV need not hold in the abstract structure is shown by interpreting Σ as the polynomial ideals of the ring $K[x_1, x_2, x_3]$. On taking $M = (x_1^2, x_2^2, x_3)$, $A = (x_1^2, x_3)$ and $B = (x_2^2, x_3)$ we find that $(A, B):M = (1)$, $(A:M, B:M) = (x_1^2, x_2^2, x_3)$. On the other hand, if $M = (x_1x_2, x_3^2)$, $A = (x_1^2, x_3^2)$, $B = (x_2^2, x_3^2)$, then $M:[A, B] = (1)$, $(M:A, M:B) = (x_1, x_2, x_3^2)$.

We now make the following assumption:

Postulate C. *If A divides B, there exists an element P such that $A = B:P$.*

The following consequences of C are needed in the discussion of (5.1) and (5.2) and serve to reveal its scope.

LEMMA 5.1. *If P and Q are any two elements of Σ, then $(P, Q) = P:(P:Q)$.*

Proof. $(P, Q) \supset P \rightarrow (P, Q) = P:N$ by Postulate C. By Theorem 4.1 and rule (4.3), $(P, Q) = P:N = P:(P:(P:N)) = P:(P:(P, Q)) = P:(P:Q)$.

LEMMA 5.2. *If $Q \supset P$, then $Q = P:(P:Q)$.*

Proof. $Q \supset P \rightarrow Q = (P, Q) \rightarrow Q = P:(P:Q)$ by Lemma 5.1.

LEMMA 5.3. *If $P \supset M$ and $Q \supset M$, then $M:P \supset M:Q$ only if $Q \supset P$.*

Proof. By Lemma 5.2, $P \supset M, Q \supset M \rightarrow P = M:(M:P), Q = M:(M:Q)$. Then by rule (4.51), $M:P \supset M:Q \rightarrow M:(M:Q) \supset M:(M:P) \rightarrow Q \supset P$.

This lemma is a limited converse of the first part of rule (4.51) which states that $Q \supset P$ implies $M:P \supset M:Q$. The direct converse is of course false.

THEOREM 5.1. *Lemma 5.1 and Postulate C are equivalent.*

Proof. Lemma 5.1 implies Lemma 5.2. And Lemma 5.2 is Postulate C with $N = P:Q$.

LEMMA 5.31. *If $P \supset M$ and $Q \supset M$, then $M:P = M:Q$ if and only if $Q = P$.*

Proof. $Q = P \rightarrow M:P = M:Q$ by (4.51). $M:P = M:Q \rightarrow Q = P$ if $P \supset M$, $Q \supset M$ by Lemma 5.3.

LEMMA 5.4. *If $M:N = M$, then $(M, N) = I$.*

Proof. By Lemma 5.1 and rule (4.2), $(M, N) = M:(M:N) = M:M = I$.

Lemma 5.4 shows that the distinction between "relatively prime" and "coprime" (Teilerfremd) elements, which must be made in the general theory (van der Waerden [2], Chapter XII, p. 30), vanishes if Postulate C is assumed.

THEOREM 5.2. *Lemma 5.3 and Postulate C are equivalent to one another.*

Proof. We have shown that Postulate C $\rightarrow$ Lemma 5.1 $\rightarrow$ Lemma 5.2 $\rightarrow$ Lemma 5.3 $\rightarrow$ Lemma 5.31; Lemma 5.2 $\rightarrow$ Postulate C. It suffices then to show that Lemma 5.31 $\rightarrow$ Lemma 5.2. By Theorem 4.1 and rule (4.21), $P:(P, Q) = P:Q = P:(P:(P:Q))$. Also $(P, Q) \supset P$, $P:(P:Q) \supset P$. Hence by Lemma 5.31, $(P, Q) = P:(P:Q)$. This is Lemma 5.2.

6. We shall now prove the important

THEOREM 6.1. *If Postulate C holds, the structure Σ is arithmetic.*[9]

Proof. It suffices to show that

(i) $$(C, [A, B]) \supset [(C, A), (C, B)];$$

for we have trivially $[(C, A), (C, B)] \supset (C, [A, B])$.

Assume Postulate C. Then by Lemma 5.1 and the first distributive law, $(C, [A, B]) = [A, B]:\{[A, B]:C\} = [A, B]:\{[A:C, B:C]\} = [A, B]:M$, where we have written M for $[A:C, B:C]$. Thus by the first distributive law, and Lemma 5.1,

$$(C, [A, B]) = [A:M, B:M],$$

$$[(C, A), (C, B)] = [A:(A:C), B:(B:C)].$$

Now $A:C \supset M$, $B:C \supset M$. Hence by rule (4.51), $A:M \supset A:(A:C)$, $B:M \supset B:(B:C)$, so that $[A:M, B:M] \supset [A:(A:C), B:(B:C)]$, giving (i).

[9] Stated by Garrett Birkhoff ([6], p. 619) for the special instance of the ideals of a commutative ring. We have made no assumption here that our structure is Dedekindian.

Postulate C need not hold in an arithmetic structure. For consider the numbers 1, 2 and 4 with both multiplication and cross-cut taken as L. C. M. and union as G. C. D. In this arithmetic structure, $1:1 = 1:2 = 1:4 = 2:2 = 2:4 = 4:4 = 1$; $2:1 = 2$; $4:1 = 4:2 = 4$. However, $4:(4:2) = 1$, $(4, 2) = 2$. This contradicts Lemma 5.1.

If Σ is a Boolean algebra and we interpret both multiplication and cross-cut as Boolean multiplication and union as Boolean addition, then $A:B = A + B'$ where B' is the negative of B. Hence

$$A:(A:B) = A:(A + B') = A + (A + B')' = A + A'B = A + B = (A, B),$$

so that Postulate C is satisfied by Theorem 5.1.

LEMMA 6.1. *If Postulate* C *holds, then for any three elements A, B and M of Σ,*

$$(M:A, M:B) = M:[(M, A), (M, B)].$$

Proof. Assume Postulate C. Then by Lemma 5.1, $[(M, A), (M, B)] = [M:(M:A), M:(M:B)] = M:(M:A, M:B)$, by the first distributive law. Now $(M:A, M:B) \supset M$. Hence by Lemma 5.2, $(M:A, M:B) = M:\{M:(M:A, M:B)\} = M:[(M, A), (M, B)]$.

THEOREM 6.2. *If Postulate* C *holds, then for any three elements A, B and M of Σ*

$$M:[A, B] = (M:A, M:B). \tag{5.2}$$

Proof. Assume Postulate C. Then by Theorem 6.1, $[(M, A), (M, B)] = (M, [A, B])$. Hence by Lemma 6.1, $(M:A, M:B) = M:(M, [A, B]) = M:[A, B]$, by rule (4.21).

7. THEOREM 7.1. *If Postulate* C *holds, then*

$$(A:B, B:A) = I.$$

Proof. Assume Postulate C. Then by rule (4.21), Theorem 6.2 and rule (4.2), $(A:B, B:A) = ([A, B]:B, [A, B]:A) = [A, B]:[A, B] = I$.

THEOREM 7.2. *If Postulate* C *holds, then for any three elements A, B and M of Σ*

$$(A, B):M = (A:M, B:M). \tag{5.1}$$

Proof. $(A, B) \supset A \rightarrow (A, B):M \supset A:M$ by (4.31). Hence $(A, B):M \supset (A:M, B:M)$ and it suffices to show that

$$(A:M, B:M) \supset (A, B):M. \tag{i}$$

$(4.2) \rightarrow A:M \supset A$, $B:M \supset B \rightarrow (A:M, B:M) \supset (A, B)$. Also $(4.2) \rightarrow (A, B):M \supset (A, B)$. Hence by Lemma 3.3, (i) follows if

$$(A, B):\{(A, B):M\} \supset (A, B):\{(A:M, B:M)\}. \tag{ii}$$

By Lemma 5.1, the left side of (ii) is $((A, B), M) = ((A, M), (B, M))$. By the first distributive law, the right side is $[(A, B):(A:M), (A, B):(B:M)]$.

Now by Lemma 5.1 and rule (4.4), $(A, B):(A:M) = \{A:(A:B)\}:(A:M) = \{A:(A:M)\}:(A:B) = (A, M):(A:B)$. Similarly, we have $(A, B):(B:M) = (B, M):(B:A)$. Thus the right side of (ii) equals $[(A, M):(A:B), (B, M):(B:A)]$. By Theorem 7.1, $(A:B, B:A) = I$. Hence by Lemma 4.1, $((A, M), (B, M)) \supset [(A, M):(A:B), (B, M):(B:A)]$. This gives (ii).

IV. Multiplicative properties of structures

8. We shall deduce here two or three interesting consequences of the fourth distributive law. These are independent of Postulate C, which we no longer assume.

THEOREM 8.1. *If the fourth distributive law holds, then multiplication is distributive with respect to cross-cut; that is, for any three elements A, B and M of Σ*

$$M[A, B] = [MA, MB]. \tag{8.1}$$

Proof. Let N be any element of Σ. Then by (5.2) and (4.4) $N:[MA, MB] = (N:MA, N:MB) = ((N:M):A, (N:M):B) = (N:M):[A, B] = N:M[A, B]$.

On taking $N = [MA, MB]$ and $N = M[A, B]$ and applying rule (4.3) we see that $M[A, B] \supset [MA, MB]$, $[MA, MB] \supset M[A, B]$. Hence (8.1) follows.

Let $P_1, Q_1; P_2, Q_2; \cdots; P_n, Q_n$ be n pairs of elements mutually residual with respect to a fixed element M, so that

$$M:P_i = Q_i, \qquad M:Q_i = P_i, \qquad (i = 1, 2, \cdots, n).$$

Then the first and fourth distributive laws show that

$$M:(Q_1, Q_2, \cdots, Q_n) = [P_1, P_2, \cdots, P_n],$$

$$M:[P_1, P_2, \cdots, P_n] = (Q_1, Q_2, \cdots, Q_n).$$

Hence we have

THEOREM 8.2. *If the fourth distributive law holds and P_i, Q_i are mutually residual with respect to M for $i = 1, 2, \cdots, n$, then $[P_1, P_2, \cdots, P_n]$ and $(Q_1, Q_2, \cdots, Q_n)$ are also mutually residual with respect to M.*

Pairs of mutually residual elements thus form a kind of structure.

LEMMA 8.1. *If multiplication is distributive with respect to cross-cut, then for any two elements A, B of Σ*

$$AB = [A, B](A, B). \tag{8.2}$$

Proof. We always have $[A, B] \supset AB \supset [A, B](A, B)$. Assume that (8.1) holds. Then $[A, B](A, B) = [A(A, B), B(A, B)]$. Since $A(A, B) \supset AB$, $B(A, B) \supset AB$, we have $[A, B](A, B) \supset AB$. This gives (8.2).

Formula (8.2) may be generalized as follows. Let $S_1, S_2, \cdots, S_n$ be any n elements of Σ, and let

$$T_1 = S_2 \cdot S_3 \cdot \cdots \cdot S_n, \quad T_2 = S_1 \cdot S_3 \cdot \cdots \cdot S_n, \cdots, T_n = S_1 \cdot S_2 \cdot \cdots \cdot S_{n-1}.$$

Then if formula (8.2) holds, we readily prove by induction that[10]

$$(8.3) \qquad S_1 \cdot S_2 \cdot \cdots \cdot S_n = [S_1, S_2, \cdots, S_n](T_1, T_2, \cdots, T_n).$$

Formulas (8.2), (8.3) are thus consequences of the fourth distributive law.

9. We conclude by giving some properties of structures which are semi-groups. The structure Σ is said to form a semi-group if

(9.1) *For any three elements A, B, C of Σ, $AB = AC$ implies $B = C$.*

This condition is easily seen to be equivalent to

(9.2) *For any three elements A, B, C of Σ, $AB \supset AC$ implies $B \supset C$.*

An equivalent condition in terms of residuation is

(9.3) *For any two elements M and N of Σ, $MN:M = N$.*

In a semi-group, a quotient $Q = \dfrac{A}{B}$ need not exist if $B \supset A$. But if it does exist, it is unique in the sense that $A = QB$ is satisfied for only one value of Q. (9.1) may be restated as

(9.4) *For any two elements M and N of Σ, there is at most one element R such that $MR = N$.*

The equivalence of (9.1), (9.2), (9.3), (9.4) is independent of Postulate C.

THEOREM 9.1. *Let Σ be a structure in which multiplication is distributive with respect to cross-cut. Then if Σ is also a semi-group, Σ is an arithmetic structure.*[11]

Proof. We are to show that (9.1) implies that

$$(2.1) \qquad (A, [B, C]) = [(A, B), (A, C)].$$

By hypothesis and (8.1), (8.2) of the previous section

$$[AB, AC] = A[B, C] = [A, [B, C]](A, [B, C]) = [[A, B], [A, C]](A, [B, C]).$$

$$\begin{aligned} &[[A, B], [A, C]][(A, B), (A, C)] \\ &\qquad = [[A, B](A, B), [A, B](A, C), [A, C](A, B), [A, C](A, C)] \\ &\qquad = [AB, AC, [A, B](A, C), [A, C](A, B)] \\ &\qquad = [[AB, AC], M], \end{aligned}$$

where

$$M = [[A, B](A, C), [A, C](A, B)].$$

But $[AB, AC] \supset [[AB, AC], M]$. Hence

$$[[A, B], [A, C]](A, [B, C]) \supset [[A, B], [A, C]][(A, B), (A, C)].$$

[10] This formula is of great importance in common arithmetic. See, for example, Stieltjes [5], Chapter I, §§7–10.

[11] The converse of this theorem is of course false as is shown by the structure consisting of the finite ring of integers modulo 4.

Therefore by (9.2) and hypothesis

$$(A, [B, C]) \supset [(A, B), (A, C)].$$

But we have trivially $[(A, B), (A, C)] \supset (A, [B, C])$. Hence (2.1) follows.

On combining this result with Theorem 8.1, we obtain

THEOREM 9.2. *In a semi-group, the fourth distributive law for residuation is a sufficient condition that the structure be arithmetic.*

REFERENCES

1. O. ORE, Annals of Math., vol. 36 (1935), pp. 406–437.
2. B. L. VAN DER WAERDEN, *Moderne Algebra*, vol. 2, Berlin, 1931.
3. F. S. MACAULAY, *The Algebraic Theory of Modular Systems*, Cambridge, 1916.
4. R. DEDEKIND, *Über die Theorie der ganzen algebraischen Zahlen, Werke*, vol. III, pp. 1–222, especially §170.
5. T. J. STIELTJES, Annales de Toulouse, (1), vol. 4 (1890), pp. 1–102.
6. GARRETT BIRKHOFF, Bull. Amer. Math. Soc., vol. 40 (1934), pp. 613–619.

CALIFORNIA INSTITUTE OF TECHNOLOGY.

ASYMPTOTIC RELATIONS FOR DERIVATIVES

By R. P. Boas, Jr.

1. **Introduction.** A well known theorem of Hardy and Littlewood states, in a form in which it is often quoted, that if $f(x)$ is of class C^2 on $(0, \infty)$, and if, as $x \to \infty$, $f(x) = o(1)$ and $f''(x) = O(x^{-2})$, then $f'(x) = o(x^{-1})$. It is a special case of a theorem in which the powers of x in the order relations are replaced by more general functions; and this in turn can be used to establish an extended theorem where from the order of $f(x)$ and of $f^{(n)}(x)$ $(n \geqq 2)$ one deduces the orders of the intermediate derivatives.[1] Now, if we think of the hypothesis on $f''(x)$ in the original theorem as "$x^2 f''(x) = O(1)$", it is a hypothesis on the order of the function resulting from applying a certain linear differential operator to $f(x)$. The principal result of this note is the corresponding theorem when the operator $x^2 \dfrac{d^2}{dx^2}$ is replaced by a certain more general, n-th order, linear operator, L; from the order of $L[f(x)]$ and of $f(x)$, the order of

$$f^{(k)}(x) \qquad (k = 1, 2, \cdots, n-1)$$

can be deduced. This result, and a preliminary theorem, overlap the results of Hardy and Littlewood, but neither include them nor are included by them. The full statement of our main theorem is somewhat complex; to illustrate it as simply as possible, a special case, sufficient for many applications, will be stated here.

Let

$$L[f(x)] = \sum_{i=0}^{n} A_i x^i f^{(i)}(x), \tag{1.1}$$

where the A_i are constants, $A_n \neq 0$. Let $f(x)$ be of class C^n on $(0, \infty)$, and suppose that as $x \to \infty$, $L[f(x)] < O(1)$. If $f(x) = O(1)$, then $f^{(k)}(x) = O(x^{-k})$; if $f(x) = o(1)$, then $f^{(k)}(x) = o(x^{-k})$ $(k = 1, 2, \cdots, n-1)$.

Examples of operators $L[f(x)]$ which have the form (1.1) are $x^n f^{(n)}(x)$; the operator

$$L_{k,x}[f(x)] = \frac{(-x)^{k-1}}{k!\,(k-2)!} \frac{d^{2k-1}}{dx^{2k-1}} [x^k f(x)] \qquad (k \geqq 2)$$

Received June 23, 1937; presented to the American Mathematical Society, September, 1937.

[1] G. H. Hardy and J. E. Littlewood, *Contributions to the arithmetic theory of series*, Proceedings of the London Mathematical Society, (2), vol. 11 (1913), pp. 411–478; 417 ff.

Reference should also be made to E. Landau, *Über einen Satz von Herrn Esclangon*, Mathematische Annalen, vol. 102 (1929–30), pp. 177–188. Landau considers more general differential operators than we do, but his results are less general in other respects.

used by D. V. Widder to invert the Stieltjes transform;[2] and the operator $L_{k,x}\{L_{k,x}[f(x)]\}$, which inverts the iterated Stieltjes transform.[3] In a later paper by D. V. Widder and the author, the results of this note will be applied to the theory of the iterated Stieltjes transform.

2. **Theorems with second derivatives.** We shall denote by "$\uparrow$" and "$\downarrow$", respectively, the classes of positive functions which are non-decreasing or non-increasing, $0 < x < \infty$. We shall also consider the class of functions $\varphi(x)$, positive on $0 < x < \infty$, and such that $\varphi(cx)/\varphi(x)$ is uniformly bounded for $0 < x < \infty$ and $\frac{1}{2} \leqq c \leqq 2$; we shall call it "$K$".[4] C^n denotes the class of functions having continuous n-th derivatives on $(0, \infty)$. With these notations, we can state

THEOREM 1A. *If $\varphi(x)$ and $\psi(x)$ satisfy any one of the conditions*

(a) $\varphi(x) \in \downarrow,\ \psi(x) \in \downarrow$;

(b) $\varphi(x) \in \uparrow,\ \psi(x) \in \uparrow$;

(c) $x^r\varphi(x) \in \uparrow,\ x^{r+2}\psi(x) \in \uparrow$, *with* $r \geqq 0$, *and* $\varphi(x) = O(x^2\psi(x))$;

then, for any $f(x) \in C^2$, as $x \to \infty$ all of the following statements are true:

A_1. $f(x) = O(\varphi(x))$ *and* $f''(x) = O(\psi(x))$ *imply* $f'(x) = O([\varphi(x)\psi(x)]^{\frac{1}{2}})$;

A_2. $f(x) = o(\varphi(x))$ *and* $f''(x) = O(\psi(x))$ *imply* $f'(x) = o([\varphi(x)\psi(x)]^{\frac{1}{2}})$;

A_3. $f(x) = O(\varphi(x))$ *and* $f''(x) = o(\psi(x))$ *imply* $f'(x) = o([\varphi(x)\psi(x)]^{\frac{1}{2}})$, *provided that* $\varphi(x) = o(x^2\psi(x))$ *is added to condition* (c).

THEOREM 1B. *If* $\varphi(x) \in K$, $\psi(x) \in K$, $\varphi(x) = O(x^2\psi(x))$, *and* $f(x) \in C^2$, *then, as* $x \to \infty$,

B_1. $f(x) = O(\varphi(x))$ *and* $f''(x) < O(\psi(x))$ *imply* $f'(x) = O([\varphi(x)\psi(x)]^{\frac{1}{2}})$;

B_2. $f(x) = o(\varphi(x))$ *and* $f''(x) < O(\psi(x))$ *imply* $f'(x) = o([\varphi(x)\psi(x)]^{\frac{1}{2}})$;

B_3. $f(x) = O(\varphi(x))$, $f''(x) < o(\psi(x))$, *and* $\varphi(x) = o(x^2\psi(x))$ *imply* $f'(x) = o([\varphi(x)\psi(x)]^{\frac{1}{2}})$.

Theorem 1A under hypotheses (b) and (c)[5] and the special case of B_2, where $\varphi(x)$ and $\psi(x)$ are powers of x,[6] are due to Hardy and Littlewood.

[2] D. V. Widder, *The Stieltjes transform*, to appear in the Transactions of the American Mathematical Society.

[3] D. V. Widder, *The iterated Stieltjes transform*, Proceedings of the National Academy of Sciences, vol. 23 (1937), pp. 242–244.

[4] Any constants a and A, $0 < a < 1 < A < \infty$, would do as well as $\frac{1}{2}$ and 2 in the definition of the class K. K contains, in particular, all those functions which J. Karamata calls "regularly increasing in the wide sense". See J. Karamata, *Sur un mode de croissance régulière des fonctions*, Mathematica (Cluj), vol. 4 (1930), pp. 38–53, 194–195; the class of functions in question is defined on p. 40, and the theorems necessary to show that it is contained in K are on pp. 40, 45.

[5] G. H. Hardy and J. E. Littlewood, op. cit., pp. 417, 425–426.

[6] G. H. Hardy and J. E. Littlewood, *Tauberian theorems concerning power series and Dirichlet's series whose coefficients are positive*, Proceedings of the London Mathematical Society, (2), vol. 13 (1914), pp. 174–191. On p. 188 they give, not precisely the theorem in question, but the corresponding theorem when $x \to 0+$.

Theorem 1B does not include Theorem 1A; for example, let $f(x) = \varphi(x) = \psi(x) = e^x$; $\varphi(x)$ and $\psi(x)$ satisfy (b) of Theorem 1A, which is applicable to $f(x)$; but there is no function $\theta(x) \epsilon K$ for which $f(x) = O(\theta(x))$. Suppose, in fact, that $e^x \leqq A\theta(x)$ $(0 < x < \infty$; A a constant), and that $\theta(cx)/\theta(x) \leqq M$ $(0 < x < \infty; c > 1)$. Then $\theta(cx) \leqq M\theta(x)$, $\theta(c^2x) \leqq M\theta(cx) \leqq M^2\theta(x)$, and by induction, $\theta(c^n x) \leqq M^n\theta(x)$ $(n = 1, 2, \cdots)$. But then $e^{c^n x} \leqq A\theta(c^n x) \leqq AM^n\theta(x)$; take $x = 1$; the result is contradictory for large n. Hence Theorem 1B cannot be applied to $f(x)$.

Theorem 1A does not include Theorem 1B, because the two-sided O-conditions of Theorem 1A cannot be replaced by one-sided O-conditions; this is shown by the example $\varphi(x) = e^x$, $\psi(x) = 1$, $f(x) = -e^x$: here $f(x) = O(\varphi(x))$, $f''(x) < 0 = o(\psi(x))$, but $f'(x) \neq O([\varphi(x)\psi(x)]^{\frac{1}{2}})$. On the other hand, if the one-sided O-conditions of Theorem 1B are replaced by two-sided O-conditions, the resulting theorem is included in Theorem 1A (c).[7] To establish this, we have only to show that it is possible to construct, given any $\theta(x) \epsilon K$, a positive function $\mu(x)$, such that for some $r \geqq 0$, $x^r\mu(x) \epsilon \uparrow$, and $\theta(x) = O(\mu(x))$, $\mu(x) = O(\theta(x))$. Since $\theta(cx)/\theta(x) \leqq M$, uniformly for $0 < x < \infty$, $\frac{1}{2} \leqq c \leqq 2$, it follows (replacing x by cx) that $\theta(c^2x)/\theta(x) \leqq M^2$, and generally that $\theta(c^n x)/\theta(x) \leqq M^n$, so that we have $\theta(cx)/\theta(x) \leqq M^n$ $(2^{-n} \leqq c \leqq 2^n;$ $n = 1, 2, \cdots)$. We assume, without loss of generality, that $\theta(0) = 0$. Choosing $r \geqq 0$ so that $M^2 2^{-r+1} < 1$, we define $\mu(x)$ by

$$\mu(x) = x^{-r} \underset{0 \leqq t \leqq x}{\text{u. b.}} (t^r\theta(t));$$

evidently $x^r\mu(x) \epsilon \uparrow$; and $\mu(x) \geqq \theta(x)$, so that $\theta(x) = O(\mu(x))$. We can define, for each x, a t_x $(0 \leqq t_x \leqq 1)$, such that $(t_x x)^r\theta(t_x x) > \frac{1}{2}x^r\mu(x)$,

$$\frac{\mu(x)}{\theta(x)} < \frac{2t_x^r\theta(t_x x)}{\theta(x)}.$$

Let $x \to \infty$. If we can show that for x sufficiently large, and for some $m \geqq 1$, $t_x > 2^{-m}$, we shall have $\mu(x)/\theta(x) < 2M^m$, and hence $\mu(x) = O(\theta(x))$; this will complete the construction. But if the t_x are not bounded from zero, we cannot have $t_x = 0$ $(x > 0)$, and there is a sequence x_k with t_{x_k} approaching zero. There are then integers $n_k \geqq 1$ with $2^{-n_k} \geqq t_{x_k} > 2^{-n_k-1}$ for sufficiently large k, and

$$\frac{\mu(x_k)}{\theta(x_k)} < \frac{M^{n_k+1}}{2^{rn_k-1}} < 1.$$

This contradicts $\mu(x) \geqq \theta(x)$.

There is no theorem corresponding to Theorem 1B with the rôles of the equality and inequality signs interchanged; that is (taking for example B_1), if $\varphi(x) \epsilon K$, $\psi(x) \epsilon K$, $\varphi(x) = O(x^2\psi(x))$, then $f(x) < O(\varphi(x))$ and $f''(x) = O(\psi(x))$ do not necessarily imply $f'(x) = O([\varphi(x)\psi(x)]^{\frac{1}{2}})$. A little reflection shows that this state of affairs is to be expected; the reader will have no difficulty in constructing examples to illustrate it.

[7] That this might be true was suggested to the author by the referee.

From the point of view of this note, Theorem 1B is more important than Theorem 1A, because it will be used in establishing our main theorem. For the sake of completeness, however, we shall give a proof of Theorem 1A under hypothesis (a) (this is the case not discussed by Hardy and Littlewood), and indicate how the same method could be used for the other cases. No originality is claimed for the proofs of Theorems 1A and 1B; the method of proof is an obvious modification of that by which the theorems are usually established when $\varphi(x) = x^k$, $\psi(x) = x^{k-2}$.[8]

Let $\delta(x)$ be any function, nowhere zero, such that $x + \delta(x) > 0$ for $x > 0$. Then Taylor's theorem with remainder of second order, applied to $f(x)$, can be written

$$f'(x) = \frac{1}{\delta(x)}\,[f(x + \delta(x)) - f(x)] - \tfrac{1}{2}\delta(x)f''(x + \theta\delta(x)), \tag{2.1}$$

where $\theta = \theta(x, \delta(x))$ satisfies $0 < \theta < 1$.

Now, if $\epsilon(x)$ is any positive function, (2.1) is valid with $\delta(x) = \epsilon(x)[\varphi(x)/\psi(x)]^{\frac{1}{2}}$, and we have

$$\frac{|f'(x)|}{[\varphi(x)\psi(x)]^{\frac{1}{2}}} \leqq \frac{1}{\epsilon(x)\varphi(x)}\,[|f(x + \delta(x))| + |f(x)|] + \frac{\epsilon(x)}{2\psi(x)}\,|f''(x + \theta\delta(x))|. \tag{2.2}$$

If we assume hypothesis (a) of Theorem 1A,

$$\varphi(x + \delta(x)) \leqq \varphi(x),$$

with a similar inequality for $\psi(x)$, and (2.2) gives

$$\frac{|f'(x)|}{[\varphi(x)\psi(x)]^{\frac{1}{2}}} \leqq \begin{cases} \epsilon(x)O(1) + O(1)/\epsilon(x), \\ \epsilon(x)O(1) + o(1)/\epsilon(x), \\ \epsilon(x)o(1) + O(1)/\epsilon(x), \end{cases} \tag{2.3}$$

according as we consider A_1, A_2, or A_3, respectively. To establish A_1, we take $\epsilon(x) \equiv 1$; to establish A_2, we take for $\epsilon(x)$ a function which is $o(1)$ but of sufficiently slow decrease; to establish A_3, we take for $\epsilon(x)$ a function of sufficiently slow increase, with $1/\epsilon(x) = o(1)$.

Theorem 1A under hypothesis (b) can be established in a similar way, but more care is needed. It is natural in this case to take $\delta(x) = -\epsilon(x)[\varphi(x)/\psi(x)]^{\frac{1}{2}}$, where $\epsilon(x)$ is still a positive function, to be chosen as before; but such an argument can be used only as $x \to \infty$ on the set E_1 of points where $\varphi(x)/\psi(x) \leqq (x/\epsilon(x))^2$, since on the complementary set, E_2, $\delta(x) < -x$, and (2.1) is no longer valid. But as $x \to \infty$ on E_2, it is legitimate to suppose that $f'(0) = 0$, and the conclusions of the theorem can then be deduced from

$$|f'(x)| \leqq \int_0^x |f''(t)|\,dt,$$

[8] See, for example, E. Landau, *Darstellung und Begründung einiger neueren Ergebnisse der Funktionentheorie*, 1929, p. 58, where a proof is given for the corresponding theorem when $x \to 0+$; the modifications necessary when $x \to \infty$ are trivial.

and the fact that $\psi(x) \in \uparrow$. Under hypothesis (c), the proof is simpler, since (c) includes an auxiliary hypothesis which in effect makes E_2 empty.

We now establish Theorem 1B. We have $\varphi(x) \leqq \sigma(x)x^2\psi(x)$, where $\sigma(x) = o(1)$ for B_3, and (without loss of generality) we may take $\sigma(x) = 1$ for B_1 and B_2. Let $\epsilon(x)$ be a positive function, $\epsilon(x) < [4\sigma(x)]^{-\frac{1}{2}}$. Let $\delta(x) = \pm\epsilon(x)[\varphi(x)/\psi(x)]^{\frac{1}{2}}$; then $|\delta(x)| < \frac{1}{2}x$, and (2.1) is valid; we may write (2.1) in the form

$$\mp f'(x) = \frac{-1}{\epsilon(x)}\left(\frac{\psi(x)}{\varphi(x)}\right)^{\frac{1}{2}}[f(x+\delta(x)) - f(x)] + \frac{\epsilon(x)}{2}\left(\frac{\varphi(x)}{\psi(x)}\right)^{\frac{1}{2}} f''(x+\theta\delta(x)), \tag{2.4}$$

where $0 < \theta < 1$, and the sign taken in (2.4) is that of $-\delta(x)$. Now, $f''(x) \leqq \eta(x)\psi(x)$, where $\eta(x) > 0$, $\eta(x) = O(1)$ for B_1 and B_2, and $\eta(x) = o(1)$ for B_3. The function $\lambda(x) = \underset{x/2 \leqq t \leqq 3x/2}{\text{u.b.}} \eta(t)$ has the same properties, and $\eta(x + \theta\delta(x)) \leqq \lambda(x)$. Then

$$\frac{\mp f'(x)}{[\varphi(x)\psi(x)]^{\frac{1}{2}}} \leqq -\frac{f(x+\delta(x)) - f(x)}{\epsilon(x)\varphi(x)} + \frac{\epsilon(x)\lambda(x)\psi(x+\theta\delta(x))}{2\psi(x)}. \tag{2.5}$$

Since $\psi(x) \in K$, $\varphi(x) \in K$, and $|\delta(x)| < \frac{1}{2}$, there is a constant B such that

$$\psi(x+\theta\delta(x))/\psi(x) \leqq B, \qquad \varphi(x+\delta(x))/\varphi(x) \leqq B; \tag{2.6}$$

and hence

$$\frac{|f(x+\delta(x))|}{\varphi(x)} = \frac{|f(x+\delta(x))|}{\varphi(x+\delta(x))}\,\frac{\varphi(x+\delta(x))}{\varphi(x)} \leqq B\frac{|f(x+\delta(x))|}{\varphi(x+\delta(x))},$$

so that

$$|f(x+\delta(x))|/\varphi(x) \leqq \tau(x), \tag{2.7}$$

where $\tau(x) = O(1)$ or $o(1)$, according as we consider B_1 and B_3, or B_2. Then (2.6) and (2.7) reduce (2.5) to

$$\mp f'(x)[\varphi(x)\psi(x)]^{-\frac{1}{2}} \leqq \tau(x)/\epsilon(x) + \tfrac{1}{2}B\,\epsilon(x)\lambda(x).$$

For B_1, we take $\epsilon(x) = \frac{1}{4}$; for B_2, we take $\epsilon(x) = \min[\frac{1}{4}, (\tau(x))^{\frac{1}{2}}]$; for B_3, we take $\epsilon(x) = \min[(\lambda(x))^{-\frac{1}{2}}, (5\sigma(x))^{-\frac{1}{2}}]$.

3. **A definition and a lemma.** It is convenient to have a name for the differential operators which we shall consider. We introduce the following

DEFINITION. *A linear differential operator*

$$L[f(x)] = \sum_{i=0}^{n} p_{n-i}(x)f^{(i)}(x) \tag{3.1}$$

with

$$p_i(x) = \sum_{j=0}^{n-i} b_{ij}x^j, \tag{3.2}$$

where the b_{ij} *are constants, and* $b_{0n} = 1$, *shall be called a generalized Euler operator.*[9]

[9] The name "generalized Euler operator" is used because in the special case, mentioned in the introduction, when $p_i(x) = b_i x^{n-i}$, $L[f(x)] = g(x)$ is what is known as an Euler differential equation.

LEMMA. *If $L[f(x)]$ is a generalized Euler operator of order n, $f(x)$ is of class C^n $(0 \leqq x < \infty)$, and a is an integer, $0 \leqq a \leqq n-1$, then as $x \to \infty$,*

$$\begin{aligned}\int_0^x (x-t)^a L[f(t)]dt = x^n f^{(n-a-1)}(x)(c+o(1)) + \sum_{i=0}^{n-a-2} f^{(i)}(x)O(x^{a+i+1}) \\ + \sum_{i=0}^{a} A_i x^i \int_0^x f(t)O(t^{a-i})\,dt + O(x^a),\end{aligned} \tag{3.3}$$

where the A_i are constants, and c is a constant $\neq 0$.[10]

Let $\bar{L}[f(x)]$ be the operator adjoint to $L[f(x)]$. Then if $g(x)$ is any function of class C^n,

$$\int_0^x g(t)L[f(t)]\,dt = \int_0^x f(t)\bar{L}[g(t)]\,dt + P[f(t), g(t)]\Big|_0^x, \tag{3.4}$$

where

$$\bar{L}[g(t)] = \sum_{i=0}^{n} (-1)^i [p_{n-i}(t)g(t)]^{(i)}, \tag{3.5}$$

$$P[f(t), g(t)] = \sum_{i=1}^{n} \sum_{j=0}^{n-i} (-1)^j f^{(n-i-j)}(t)[p_{i-1}(t)g(t)]^{(j)}.\text{[11]} \tag{3.6}$$

If $g(t) = (x-t)^a$, and the differentiations are carried out in (3.5) and (3.6), we find

$$\begin{aligned}\int_0^x (x-t)^a L[f(t)]\,dt = \sum_{i=0}^{a} A_i x^i \int_0^x f(t)P_{a-i}(t)\,dt + \sum_{i=0}^{a} A_i' x^i \\ + \sum_{i=0}^{n-a-2} f^{(i)}(x)P_{i+a+1}(x) + f^{(n-a-1)}(x)(cx^n + P_{n-1}(x)),\end{aligned}$$

where A_k and A_k' denote various constants, and $P_k(t)$ is a polynomial of degree k at most (not necessarily the same at each appearance); from this (3.3) follows at once. The details are left to the reader.

4. A theorem with generalized Euler operators. We shall consider a pair of positive functions $\theta(x)$ and $\varphi(x)$ satisfying what we shall call

CONDITIONS A.[12]

(i) $\varphi(x) \in K$;

(ii) $x\theta(x) = O(\varphi(x))$;

(iii) $\displaystyle\int_0^x \theta(t)\,dt = O(\varphi(x))$;

(iv) $\displaystyle\int_0^x \varphi(t)\,dt = O(x\varphi(x))$.[13]

[10] Actually, $c = (-1)^a a!$.

[11] See, for example, E. L. Ince, *Ordinary Differential Equations*, 1927, pp. 123-124.

[12] The author is indebted to the referee for the elimination of a redundancy in Conditions A.

[13] It is supposed throughout this section that $x \to \infty$. In particular, if $\theta(x) > 0$, and $\theta(x)$ is (in Karamata's terminology) regularly increasing in the wide sense, then $\theta(x)$ and $\varphi(x) = x\theta(x)$ satisfy Conditions A. The results needed for verifying this can be found in Karamata's paper cited in footnote 4.

We now establish our main theorem.

THEOREM 2. *Let $L[f(x)]$ be a generalized Euler operator of order n, and let $f(x)$ of class C^n on $(0, \infty)$ be a solution of the differential equation $L[f(x)] = g(x)$. Let $\theta(x)$ and $\varphi(x)$ satisfy Conditions* A. *If, as $x \to \infty$,*

$$f(x) = O(\theta(x)), \tag{4.1}$$

$$\int_0^x g(t)\,dt < O(\varphi(x)), \tag{4.2}$$

then

$$f^{(p)}(x) = O(x^{-p-1}\varphi(x)) \tag{4.3}$$

for $p = 1, 2, \cdots, n - 2$; and also for $p = n - 1$ if in addition

$$g(x) < O(\theta(x)). \tag{4.4}$$

If also

$$f(x) = o(\theta(x)) \tag{4.5}$$

and

$$\int_0^\infty \theta(t)\,dt = \infty, \tag{4.6}$$

then

$$f^{(p)}(x) = o(x^{-p-1}\varphi(x)) \tag{4.7}$$

for $p = 1, 2, \cdots, n - 2$, and also for $p = n - 1$ if (4.4) is satisfied.

If the conclusion of the second part of the theorem is needed only for $1 \leqq p \leqq n - 3$ (or $n - 2$ if (4.4) is satisfied), it is a direct consequence of the first part (even without (4.6)), by successive applications of Theorem 1B, since by Condition A(ii), $\theta(x) = O(x^{-1}\varphi(x))$, while $(x + 1)^k\varphi(x) \in K$ for any real number k if $\varphi(x) \in K$.

Theorem 2 has content only when $n \geqq 2$ if (4.4) is satisfied, and only if $n \geqq 3$ when (4.2) alone is satisfied. If $n = 2$, $L[f(x)] = x^2f''(x)$, and $x\theta(x) = \varphi(x)$, Theorem 2 reduces to the special case of B_1 and B_2 of Theorem 1B where $\varphi(x) = x^2\psi(x)$ (except that Conditions A(ii), (iii), (iv), and (4.6) are then redundant). If $L[f(x)] = x^nf^{(n)}(x)$, $\varphi(x) = x\theta(x) = x^k$, and the one-sided O-condition (4.4) is replaced by a two-sided O-condition, Theorem 2 is a special case of the extension of Theorem 1A given by Hardy and Littlewood (who use a more general function $\varphi(x)$).[14]

If we take $\theta(x) = (x + 1)^{-1} \log (x + 1)$, $\varphi(x) = [\log (x + 1)]^2$, we obtain an example illustrating the theorem when $\varphi(x) \neq x\theta(x)$; and the theorem with these functions has applications.[15]

[14] Reference in footnote 5.

[15] It was used in the author's Harvard thesis (unpublished).

To establish Theorem 2, we start from the familiar formula

$$G_m(x) \equiv \int_0^x (x-t)^m g(t)\,dt = m! \int_0^x dt_m \int_0^{t_m} dt_{m-1} \cdots \int_0^{t_2} dt_1 \int_0^{t_1} g(t)\,dt \quad (m > 0).$$

From Conditions A(ii) and (iv) we obtain, for $p \geqq 1$,

$$(4.8) \qquad \begin{aligned} \int_0^x \varphi(t) t^p\,dt &= O\left(x^p \int_0^x \varphi(t)\,dt\right) = O(x^{p+1}\varphi(x)); \\ \int_0^x \theta(t) t^p\,dt &= O\left(\int_0^x \varphi(t) t^{p-1}\,dt\right) = O(x^p \varphi(x)); \end{aligned}$$

and it follows, by use of (4.2), that

$$(4.9) \qquad G_m(x) < O(x^m \varphi(x)) \qquad (m = 0, 1, 2, \cdots).$$

Now, by Condition A(iii),

$$\int_0^x \theta(t)\,dt = O(\varphi(x)),$$

$$1/\varphi(x) = O\left\{1 \Big/ \int_0^x \theta(t)\,dt\right\} = O(1),$$

$$(4.10) \qquad x^k = O(x^k \varphi(x)) \qquad (k \geqq 0).$$

Since $f(x) = O(\theta(x))$, from Condition A(iii) and (4.8), we obtain

$$(4.11) \qquad \sum_{i=0}^{a} A_i x^i \int_0^x f(t) O(t^{a-i})\,dt = O(x^a \varphi(x)).$$

We use (4.10) and (4.11) in (3.3), and, remembering that $L[f(x)] = g(x)$, we obtain

$$(4.12) \qquad \begin{aligned} G_a(x) = x^n f^{(n-a-1)}(x)(c + o(1)) + \sum_{i=0}^{n-a-2} f^{(i)}(x) O(x^{a+i+1}) \\ + O(x^a \varphi(x)) \qquad (1 \leqq a \leqq n-1). \end{aligned}$$

Suppose now that (4.3) has been established for $1 \leqq p \leqq q$, $q \leqq n-2$. Then (4.12) for $a = n - q - 1$ gives

$$(4.13) \qquad G_{n-q-1}(x) = O(x^{n-q-1}\varphi(x)),$$

and since $G_m''(x) = m(m-1)G_{m-2}(x)$ ($m \geqq 1$; we set $G_{-1}(x) = g(x)$), we have by (4.9) with $m = n - q - 3$,

$$(4.14) \qquad G_{n-q-1}''(x) < O(x^{n-q-3}\varphi(x))$$

when $q \leqq n-3$, and also when $q = n-2$ if (4.4) is satisfied. But since $\varphi(x) \in K$, $(x+1)^{n-q-1}\varphi(x) \in K$; and by part B_1 of Theorem 1B, (4.13) and (4.14) imply

$$(4.15) \qquad G_{n-q-1}'(x) = (n-q-1)G_{n-q-2}(x) = O(x^{n-q-2}\varphi(x)).$$

By (4.12) for $a = n - q - 2$, and by (4.3), assumed established for $1 \leqq p \leqq q$,

$$G_{n-q-2}(x) = x^n f^{(q+1)}(x)(c + o(1)) + O(x^{n-q-2}\varphi(x)). \tag{4.16}$$

Comparing (4.15) and (4.16), we have, since $c \neq 0$, (4.3) for $p = q + 1$, provided only that $1 \leqq q \leqq n - 3$, or $1 \leqq q \leqq n - 2$, according as (4.4) is not or is satisfied. Hence the first part of Theorem 2 is true as soon as we verify (4.3) for $p = 1$. To do this, we have, from (4.12) with $a = n - 1$,

$$\begin{aligned} G_{n-1}(x) &= x^n f(x)(c + o(1)) + O(x^{n-1}\varphi(x)) \\ &= O(x^{n-1}\varphi(x)) \end{aligned}$$

by use of (4.1) and Condition A(ii). Then by (4.9) with $m = n - 3$, or by (4.4) if $n = 2$,

$$G_{n-3}(x) < O(x^{n-3}\varphi(x)),$$

and by Theorem 1B,

$$\begin{aligned} G_{n-2}(x) &= x^n f'(x)(c + o(1)) + f(x)O(x^{n-1}) + O(x^{n-2}\varphi(x)) \\ &= O(x^{n-2}\varphi(x)), \end{aligned}$$

so that

$$f'(x) = O(x^{-2}\varphi(x)).$$

This completes the proof of the first part of Theorem 2. The proof of the second part is similar.

Since $\int_0^\infty \theta(t)\, dt$ diverges, in place of (4.10) we have

$$x^k = o(x^k\varphi(x)) \qquad (k \geqq 0). \tag{4.17}$$

In place of (4.11),

$$\sum_{i=0}^{a} A_i x^i \int_0^x f(t)O(t^{a-i})\, dt = o(x^a\varphi(x)).$$

Hence (3.3) gives us, in place of (4.12),

$$\begin{aligned} G_a(x) = x^n f^{(n-a-1)}(x)(c + o(1)) + \sum_{i=0}^{n-a-2} f^{(i)}(x)O(x^{a+i+1}) \\ + o(x^a\varphi(x)) \qquad (1 \leqq a \leqq n - 1). \end{aligned} \tag{4.18}$$

The remainder of the proof is exactly parallel to the proof of the first part of the theorem, part B_2 of Theorem 1B being used instead of part B_1. The details are left to the reader.

5. **Conclusion.** We have stated our theorems with $x \to \infty$. It is clear that they hold, with obvious modifications, when $x \to 0+$. The proofs are given most simply by modifying the reasoning directly, rather than by a change of

variable. The classes $\uparrow$ and $\downarrow$ are replaced by the classes of functions non-decreasing or non-increasing as $x \to 0+$; that is, $\uparrow$ becomes $\downarrow$ and reciprocally. The class K is replaced by the class of functions $\varphi(x)$ with $\varphi(1/x) \in K$. Since it may not be quite evident what becomes of Theorem 2, we state the modified theorem in detail.

THEOREM 3. *Let*

$$L[f(x)] = \sum_{i=0}^{n} p_{n-i}(x) f^{(i)}(x),$$

$$p_i(x) = \sum_{j=0}^{i} b_{ij} x^{-j}, \qquad b_{00} = 1.$$

Let $f(x)$ of class C^n on $0 < x < \infty$ be a solution of the differential equation $L[f(x)] = g(x)$. Let $\theta(1/x)$ and $\varphi(1/x)$ satisfy Conditions A. If, as $x \to 0+$,

$$f(x) = O(\theta(x)),$$

$$\int_x^\infty g(t)\,dt < O(x^{-n+2}\varphi(x)),$$

then

$$f^{(p)}(x) = O(x^{-p+1}\varphi(x))$$

for $p = 1, 2, \cdots, n-2$; and also for $p = n-1$ if in addition

$$g(x) < O(x^{-n}\theta(x)). \tag{5.1}$$

If also

$$f(x) = o(\theta(x))$$

and

$$\int_0^\infty \theta(t^{-1})\,dt = \infty,$$

then

$$f^{(p)}(x) = o(x^{-p+1}\varphi(x))$$

for $p = 1, 2, \cdots, n-2$, and also for $p = n-1$ if (5.1) is satisfied.

HARVARD UNIVERSITY.

TAUBERIAN THEOREMS RELATED TO BOREL AND ABEL SUMMABILITY

BY MORRIS L. KALES

The following work is divided into two sections. The first deals with a Tauberian theorem of summability methods related to Borel summability, and the second with a Tauberian theorem of summability methods related to Abel summability.

In 1925 Robert Schmidt gave a proof of the following theorem **[6]**:[1]

If a series is Abel summable to the value s, and if the partial sums s_n satisfy the condition

$$\underline{\lim}\,(s_m - s_n) \geqq 0 \quad \textit{whenever} \quad \frac{m-n}{n} \to 0 \qquad (m > n),$$

then

$$\lim_{n\to\infty} s_n = s.$$

In the same year Schmidt gave a proof of an analogous theorem concerning Borel summability. This theorem states **[7]**:

If a series is Borel summable to the value s, and if the partial sums s_n satisfy the condition

$$\underline{\lim}\,(s_m - s_n) \geqq 0 \quad \textit{whenever} \quad \frac{m-n}{n^{\frac{1}{2}}} \to 0 \qquad (m > n),$$

then

$$\lim s_n = s.$$

In the following two years Vijayaraghavan **[10, 11]** gave a new and more elementary proof for each of these theorems.

It will be observed that in each of the preceding cases the method of summability is a power series method, and that the condition imposed on the sequence of partial sums is of the following type:

$$\text{(i)} \qquad \underline{\lim}\,(s_m - s_n) \geqq 0 \quad \text{whenever} \quad \frac{m-n}{\varphi(n)} \to 0 \qquad (m > n),$$

where $\varphi(n)$ is an increasing function of n which tends to ∞ with n.

Received June 30, 1937. The author is indebted to Professors J. D. Tamarkin and O. Szász, who suggested this problem to him.

[1] The numbers refer to the list of references at the end of this paper. For an extensive bibliography on the subject of Tauberian theorems see N. Wiener, *Tauberian theorems*, Annals of Mathematics, (2), vol. 33 (1932), pp. 1-100.

These facts suggest an interesting possibility of extension. It is natural to ask: corresponding to any increasing function $\varphi(n)$ which tends to infinity with n, does there exist a power series method of summability such that every series which is summable by that method, and whose partial sums satisfy condition (i), is convergent to the same value s? In particular we may take the case when $\varphi(n) = n^{\alpha}$ $(0 < \alpha < 1)$. [Cf. J. M. Hyslop, **2**.] This is a special case of Theorem I which is proved in the first part of this paper. When $\alpha = \frac{1}{2}$, we obtain the case of Borel summability as a special case, but the case of Abel summability, for which $\alpha = 1$, is not included in Theorem I. The latter case is the point of departure for Theorem II, which is proved in the second part of this paper.

I

By combining the methods of Vijayaraghavan and Valiron [**9**] I have been able to prove the following theorem which I now proceed to formulate.

Let

$$\sum_{1}^{\infty} g(n)e^{-G(n)}x^n = F(x) \qquad (x > 0)$$

be a power series with radius of convergence R. (R may be finite or infinite.) Let the function $G(x)$ satisfy the following conditions:

I. $G(x)$ has a second derivative $G''(x)$ which is positive and which tends monotonically to zero as x tends to infinity.

II. There exists an increasing function $\psi(x)$, which tends to infinity as $x \to \infty$, such that

$$\frac{G''(x_1)}{G''(x)} = 1 + o\left\{\frac{1}{\psi(x)}\right\} \quad \text{whenever} \quad |x_1 - x| \leqq \sqrt{\frac{\psi(x)}{G''(x)}}\,.$$

III. There exists a decreasing function $H(x) \sim G''(x)$ $(x \to \infty)$ such that $2/H(x)$ has an inverse $K(x)$ which has a continuous second derivative and satisfies for all large x relations of the form

(i) $AK(x) < xK'(x) < BK(x)$,

(ii) $x^2 \,|\, K''(x) \,|\, < CK(x)$,

where A, B, C are positive constants.

IV. $$\lim_{x\to\infty} \frac{x^2 G''(x)}{\log x} = \infty.$$

Finally, let the function $g(x)$ be defined as follows:

V. $$g(x) = x^{\sigma}L(x) > 0,$$

where σ is any real number and $L(x)$ satisfies the condition

$$\lim_{x\to\infty} \frac{L(\lambda x)}{L(x)} = 1 \quad \text{for every fixed } \lambda > 0.$$

THEOREM I. *Let the functions $G(x)$ and $g(x)$ satisfy relations* I–V *inclusive. If*

$$\lim_{x \to R} \frac{1}{F(x)} \sum_1^\infty s_n g(n) e^{-G(n)} x^n = s \qquad (s \text{ finite}),$$

and if

$$\underline{\lim}\, (s_m - s_n) \geqq 0 \text{ whenever } (m - n)\sqrt{G''(n)} \to 0 \qquad (m > n),$$

then

$$\lim_{x \to \infty} s_n = s.$$

For the cases when $g(x) \sim 1$ as $x \to \infty$, or $g(x) = x^\sigma L(x)$ with $\sigma > 0$, condition IV *is not required.*

The proof of this theorem depends upon a number of lemmas which are analogous to results obtained by Valiron and Vijayaraghavan. Before turning to these lemmas, it will be useful to enumerate a few of the immediate consequences of conditions I–V which can be easily verified.

(A) Condition II implies that

$$\lim_{x \to \infty} x^2 G''(x) = \infty.$$

(B) Without any additional assumption, the function $\psi(x)$ of condition II may be replaced by any function $\varphi(x)$ which tends to infinity as $x \to \infty$ and satisfies $\varphi(x) \leqq \psi(x)$.

(C) The series

$$\sum_1^\infty e^{-G(n)} x^n = F_1(x), \qquad \sum_1^\infty g(n) e^{-G(n)} x^n = F(x)$$

have the same radius of convergence.

(D_a) If $g(x) = x^\sigma L(x)$ with $\sigma > 0$, or $g(x) \sim 1$ as $x \to \infty$, then there exists a positive constant K such that

$$\frac{g(n)}{g(m)} < K \quad \text{if} \quad 1 \leqq n \leqq m.$$

(D_b) If $g(x) = x^\sigma L(x)$, where σ is any real number, there exist positive constants K and α such that

$$\frac{g(n)}{g(m)} < K m^\alpha \quad \text{if} \quad 1 \leqq n \leqq m.$$

(D_c) If $g(x) = x^\sigma L(x)$, where σ is any real number, there exist positive constants K and α such that

$$\frac{g(m)}{g(n)} < K\left(\frac{m}{n}\right)^\alpha \quad \text{if} \quad 1 \leqq n \leqq m.$$

In proving C and D use is made of the following representation of the function $L(x)$ which is due to Karamata [**3**, p. 45]:

$$L(x) = c(x) \exp \int_a^x t^{-1}\epsilon(t)\, dt, \tag{1}$$

where

$$\lim_{x\to\infty} c(x) = c > 0 \tag{2}$$

and

$$\lim_{x\to\infty} \epsilon(x) = 0. \tag{3}$$

Also, from (1), (2), and (3) it follows that

$$\lim_{x\to\infty} \frac{L(\lambda x)}{L(x)} = 1 \tag{4}$$

uniformly for all λ in any finite interval not containing the origin.

Let $\xi = [y + 1]$, where $y = y(x)$ is the solution of the equation

$$G'(y) = \log x.$$

We are now ready to prove

LEMMA 1. *Let $N_1 = N_1(\xi)$ and $N_2 = N_2(\xi)$ be two positive integers which satisfy the relations*

$$\text{(i)} \qquad N_2 - \xi = \xi - N_1 \qquad (N_2 > \xi),$$

$$\text{(ii)} \qquad \lim_{\xi\to\infty} (N_2 - \xi)^2 G''(N_2) = \infty.$$

Then

$$\lim_{x\to R} \frac{1}{F(x)} \sum_{N_1+1}^{N_2-1} g(n)e^{-G(n)}x^n = 1.$$

Lemma 1 is an immediate consequence of Lemmas 1_a and 1_b which follow.

LEMMA 1_a.

$$\lim_{x\to R} \frac{1}{F(x)} \sum_{1}^{N_1} g(n)e^{-G(n)}x^n = 0.$$

Let M be any positive integer such that $1 \leqq M \leqq \xi - 1$. Writing $T_n(x)$ for $e^{-G(n)}x^n$ and following the method of Valiron [**9**], we can easily show that

$$\frac{T_{M-p}}{T_{\xi-1-p}} \leqq e^{-\frac{1}{2}(M-\xi+1)^2G''(\xi)} \qquad (0 \leqq p \leqq M - 1). \tag{1}$$

To prove this, we first apply Taylor's expansion and get

$$G(M - p) = G(\xi - 1 - p) + (M - \xi + 1)G'(\xi - 1 - p) + \frac{(M - \xi + 1)^2}{2} G''(M_p), \tag{2}$$

where $M - p < M_p < \xi - 1 - p$, and then make use of the facts that $G''(x)$ is a decreasing function of x, that $G'(x)$ is an increasing function of x, and that

$$\log x = G'(y) \geqq G'(\xi - 1) \geqq G'(\xi - 1 - p).$$

Now let us consider the case where $g(x) = x^\sigma L(x)$ with $\sigma > 0$, or $g(x) \sim 1$ as $x \to \infty$. Then by D_a

$$\frac{g(M - p)}{g(\xi - 1 - p)} < K. \tag{3}$$

Combining (1) and (3) and placing $M = N_1$ we get

$$\frac{g(N_1 - p)T_{N_1-p}}{g(\xi - 1 - p)T_{\xi-1-p}} < Ke^{-\frac{1}{2}(N_1-\xi+1)^2G''(\xi)} \qquad (0 \leqq p \leqq N_1 - 1). \tag{4}$$

Hence

$$\begin{aligned} \sum_{n=1}^{N_1} g(n)T_n(x) &= \sum_{p=0}^{N_1-1} g(N_1 - p)T_{N_1-p} \\ &< Ke^{-\frac{1}{2}(N_1-\xi+1)^2G''(\xi)} \sum_{p=0}^{N_1-1} g(\xi - 1 - p)T_{\xi-1-p} < Ke^{-\frac{1}{2}(N_1-\xi+1)^2G''(\xi)}\cdot F(x). \end{aligned} \tag{5}$$

Since $\xi < N_2$, and therefore $G''(\xi) \geqq G''(N_2)$, it follows from (i) and (ii) that

$$\lim_{\xi\to\infty} (N_1 - \xi + 1)^2 G''(\xi) = \infty. \tag{6}$$

Combining (5) and (6) we have

$$\lim_{x\to R} \frac{1}{F(x)} \sum_{1}^{N_1} g(n)T_n(x) = 0. \tag{7}$$

Now let us consider the case where $g(x) = x^\sigma L(x)$ and σ is any real number. Then by D_b positive constants K and α exist such that

$$\frac{g(M - p)}{g(\xi - 1 - p)} < K(\xi - 1 - p)^\alpha < K\xi^\alpha \tag{8}$$
$$(1 \leqq M \leqq \xi - 1;\ p = 0, 1, 2, \cdots, M - 1).$$

Combining (1) and (8) we have

$$\frac{g(M - p)T_{M-p}}{g(\xi - 1 - p)T_{\xi-1-p}} < K\xi^\alpha e^{-\frac{1}{2}(M-\xi+1)^2G''(\xi)}. \tag{9}$$

Let δ be a fixed positive number such that $N = [(1 - \delta)\xi] + 1$. Then writing

$$\frac{1}{F(x)} \sum_{1}^{N_1} g(n)T_n(x) = \frac{1}{F(x)} \sum_{1}^{N} g(n)T_n + \frac{1}{F(x)} \sum_{N+1}^{N_1} g(n)T_n = I_1 + I_2, \tag{10}$$

we obtain from (9)

$$I_1 < K\xi^\alpha e^{-\frac{1}{2}(N-\xi+1)^2G''(\xi)} \leqq Ke^{-\frac{1}{2}\delta^2\xi^2G''(\xi)+\alpha\log\xi}. \tag{11}$$

But by condition IV and A

$$\lim_{\xi\to\infty} \tfrac{1}{2}\delta^2\xi^2 G''(\xi) - \alpha\log\xi = \lim \xi^2 G''(\xi)\left\{\frac{\delta^2}{2} - \frac{\alpha\log\xi}{\xi^2 G''(\xi)}\right\} = \infty. \tag{12}$$

Thus, combining (11) and (12), we get

$$\lim_{x\to R} I_1 = 0. \tag{13}$$

Since $\lim_{x\to\infty} L(\lambda x)/L(x) = 1$ uniformly for λ in any finite interval not containing the origin, it follows that, for sufficiently large ξ, if $(1-\delta)\xi < n < m < \xi$, then

$$\frac{g(n)}{g(m)} = \frac{n^\sigma L(n)}{m^\sigma L(m)} < 2\left(\frac{1}{1-\delta}\right)^{|\sigma|} = K. \tag{14}$$

Setting $M = N_1$ in (1) and combining with (14) we see that

$$\frac{g(N_1-p)T_{N_1-p}}{g(\xi-1-p)T_{\xi-1-p}} < Ke^{-\frac{1}{2}(N_1-\xi+1)^2G''(\xi)} \qquad (0 \leqq p < N_1-(1-\delta)\xi). \tag{15}$$

Hence

$$\begin{aligned} I_2 = \frac{1}{F(x)}\sum_{p=0}^{N_1-N-1} g(N_1-p)T_{N_1-p} &< K\exp\left[-\tfrac{1}{2}(N_1-\xi+1)^2G''(\xi)\right] \\ \cdot \frac{1}{F(x)}\sum_{p=0}^{N_1-N-1} g(\xi-1-p)T_{\xi-1-p} &< K\exp\left[-\tfrac{1}{2}(N_1-\xi+1)^2G''(\xi)\right]. \end{aligned} \tag{16}$$

From (16) we see that

$$\lim_{x\to R} I_2 = 0. \tag{17}$$

Combining (13) and (17) gives

$$\lim_{x\to R}\frac{1}{F(x)}\sum_{1}^{N_1} g(n)T_n(x) = 0. \tag{18}$$

It will be observed that in proving Lemma 1_a, condition IV was not required for the case where $g(x) = x^\sigma L(x)$ with $\sigma > 0$, or $g(x) \sim 1$. Since this is the only place where condition IV is used, it follows that condition IV is not required for the theorem if $g(x) = x^\sigma L(x)$ with $\sigma > 0$, or $g(x) \sim 1$.

LEMMA 1_b.

$$\lim_{x\to R}\frac{1}{F(x)}\sum_{N_2}^{\infty} g(m)e^{-G(m)}x^m = 0.$$

If we apply the theorem of the mean twice, it is easy to show that

$$\frac{T_m}{T_n} \leqq e^{-(m-n)(n-\xi)G''(n)} \qquad (\xi \leqq n \leqq m). \tag{1}$$

In particular it follows from (1) that $T_m/T_n \leqq 1$; i.e., the terms $T_n(x)$ are monotone decreasing with respect to n for $n \geqq \xi$.

By D_c there exist positive constants K and α such that

$$\frac{g(m)}{g(n)} < K\left(\frac{m}{n}\right)^{\alpha} \qquad (1 \leqq n \leqq m). \tag{2}$$

Hence

$$\frac{g(m)T_m}{g(n)T_n} < K\left(\frac{m}{n}\right)^{\alpha} e^{-(m-n)(n-\xi)G''(n)}. \tag{3}$$

Thus

$$\begin{aligned} \frac{1}{F(x)} \sum_{m=n}^{\infty} g(m)T_m &= \frac{g(n)T_n}{F(x)} \sum_{m=n}^{\infty} \frac{g(m)T_m}{g(n)T_n} \\ &< \frac{Kg(n)T_n}{F(x)} \sum_{m=n}^{\infty} \left(\frac{m}{n}\right)^{\alpha} e^{-(m-n)(n-\xi)G''(n)}. \end{aligned} \tag{4}$$

But

$$\begin{aligned} \sum_{m=n}^{\infty} \left(\frac{m}{n}\right)^{\alpha} e^{-(m-n)(n-\xi)G''(n)} &\leqq 2^{\alpha} \sum_{m=n}^{\infty} \left\{1 + \frac{(m-n)^{\alpha}}{n^{\alpha}}\right\} e^{-(m-n)(n-\xi)G''(n)} \\ &\leqq K_{\alpha}\left\{1 + \frac{1}{(n-\xi)G''(n)} + \frac{1}{(n-\xi)^{1+2\alpha}G''(n)^{1+\alpha}}\right\}. \end{aligned} \tag{5}$$

Combining (4) and (5) we have

$$\frac{1}{F(x)} \sum_{m=n}^{\infty} g(m)T_m < \frac{KK_{\alpha}g(n)T_n}{F(x)}\left\{1 + \frac{1}{(n-\xi)G''(n)} + \frac{1}{(n-\xi)^{1+2\alpha}G''(n)^{1+\alpha}}\right\}. \tag{6}$$

Now by (2) we have, if $\xi \leqq n \leqq m$,

$$\frac{g(n)T_n}{g(m)T_m} > \frac{1}{K}\left(\frac{n}{m}\right)^{\alpha}. \tag{7}$$

Let $N = [(1 + \delta)\xi]$, where δ is a fixed positive number such that $0 < \delta < 1$, and let M be any positive integer such that $\xi \leqq M \leqq N$. Then

$$\begin{aligned} \sum_{n=\xi}^{M} g(n)T_n = g(M)T_M \sum_{n=\xi}^{M} \frac{g(n)T_n}{g(M)T_M} &> g(M)T_M \sum_{n=\xi}^{M} \frac{1}{K}\left(\frac{n}{M}\right)^{\alpha} \\ &> \frac{g(M)T_M}{K}(M - \xi)\left(\frac{1}{1-\delta}\right)^{\alpha}. \end{aligned} \tag{8}$$

Thus

$$\frac{g(M)T_M}{F(x)} < \sum_{n=\xi}^{M} \frac{g(n)T_n}{F(x)} \cdot \frac{K(1+\delta)^{\alpha}}{M - \xi} < \frac{K(1+\delta)^{\alpha}}{M-\xi}. \tag{9}$$

Let $K^2K_{\alpha}(1 + \delta)^{\alpha} = K_1$. Then if $N \leqq N_2$, we set $n = N$ in (6) and $M = N$ in (9), and combining the two we get

$$\begin{aligned} \frac{1}{F(x)} \sum_{m=N_2}^{\infty} g(m)T_m &< \frac{1}{F(x)} \sum_{m=N}^{\infty} g(m)T_m \\ &< K_1\left\{\frac{1}{N-\xi} + \frac{1}{(N-\xi)^2G''(N)} + \frac{1}{\{(N-\xi)^2G''(N)\}^{1+\alpha}}\right\} \quad (N \leqq N_2). \end{aligned} \tag{10}$$

If $N_2 \leqq N$, we may set $M = N_2$ in (9) and $n = N_2$ in (6), and combining we obtain

$$\begin{aligned}(11)\quad \frac{1}{F(x)} \sum_{m=N_2}^{\infty} g(m)T_m < K_1 \Big\{ & \frac{1}{N_2 - \xi} + \frac{1}{(N_2 - \xi)^2 G''(N_2)} \\ & + \frac{1}{\{(N_2 - \xi)^2 G''(N_2)\}^{1+\alpha}} \Big\} \quad (N_2 \leqq N).\end{aligned}$$

Since by A, $\lim_{x \to \infty} x^2 G''(x) = \infty$, it follows that

$$(12)\quad \lim_{\xi \to \infty} \frac{1}{(N - \xi)^2 G''(N)} \leqq \lim_{N \to \infty} \frac{1}{\delta^2 N^2 G''(N)} = 0.$$

Thus from (12) and condition (ii) of Lemma 1, it follows that the right sides of (10) and (11) tend to zero as $\xi \to \infty$. Hence we conclude

$$(13)\quad \lim_{x \to R} \frac{1}{F(x)} \sum_{m=N_2}^{\infty} g(m)\, T_m = 0.$$

LEMMA 2. *Under the conditions of Lemma* 1,

$$\lim_{x \to R} \frac{\sqrt{G''(N_2)}}{F(x)} \sum_{m=N_2}^{\infty} (m - N_2) g(m) e^{-G(m)} x^m = 0.$$

By (3) in Lemma 1 we have

$$\begin{aligned}\frac{1}{F(x)} \sum_{m=n}^{\infty} (m - n) g(m) T_m &= \frac{g(n)T_n}{F(x)} \sum_{m=n}^{\infty} (m - n) \frac{g(m)T_m}{g(n)T_n} \\ (1)\quad &< \frac{Kg(n)T_n}{F(x)} \sum_{m=n}^{\infty} (m - n) \left(\frac{m}{n}\right)^{\alpha} e^{-(m-n)(n-\xi)G''(n)} \\ &< \frac{KK_\alpha g(n)T_n}{F(x)} \left\{ \frac{1}{\{(n - \xi)G''(n)\}^2} + \frac{1}{(n - \xi)^{2+2\alpha} G''(n)^{2+\alpha}} \right\}.\end{aligned}$$

Let N be defined as in Lemma 1_b. Then if $N \leqq N_2$, we have

$$\begin{aligned}(2)\quad \frac{\sqrt{G''(N_2)}}{F(x)} \sum_{m=N_2}^{\infty} (m - N_2) g(m) T_m &\leqq \frac{\sqrt{G''(N)}}{F(x)} \sum_{m=N}^{\infty} (m - N) g(m) T_m \\ &< K_1 \left\{ \frac{1}{(N - \xi)^2 G''(N)} + \frac{1}{\{(N - \xi)^2 G''(N)\}^{1+\alpha}} \right\},\end{aligned}$$

and if $N_2 \leqq N$

$$\begin{aligned}(3)\quad & \frac{\sqrt{G''(N_2)}}{F(x)} \sum_{m=N_2}^{\infty} (m - N_2) g(m) T_m \\ & \qquad < K_1 \left\{ \frac{1}{(N_2 - \xi)^2 G''(N_2)} + \frac{1}{\{(N_2 - \xi)^2 G''(N_2)\}^{1+\alpha}} \right\}.\end{aligned}$$

From (2) and (3) we conclude, as in Lemma 1_b, that

$$\lim_{x\to R}\frac{\sqrt{G''(N_2)}}{F(x)}\sum_{m=N_2}^{\infty}(m-N_2)g(m)T_m=0. \tag{4}$$

The following lemma concerning the slowly decreasing sequence $\{s_n\}$ can be proved in precisely the same manner as the analogous lemma proved by Vijayaraghavan [**11**, p. 319, Lemma ϵ]:

LEMMA 3. *If* $\underline{\lim}\,(s_m - s_n) \geqq 0$ *whenever* $(m-n)\sqrt{G''(n)} \to 0$ $(m > n)$, *then to every positive number* c *there corresponds a positive number* $k = k(c)$ *such that*

$$s_m - s_n > -\{k(m-n)\sqrt{G''(n)} + c\}.$$

LEMMA 4. *Under the conditions of Theorem* I, *if*

$$\lim_{x\to R}\frac{1}{F(x)}\sum_{1}^{\infty}s_n g(n)e^{-G(n)}x^n = s \qquad (s\ \textit{finite}),$$

and if $\underline{\lim}\,(s_m - s_n) \geqq 0$ *whenever* $(m-n)\sqrt{G''(n)} \to 0$ $(m>n)$, *then*

$$s_n = O(1).$$

With the aid of Lemmas 1, 2, and 3 this lemma can be proved in the same way as Lemma 1 of Vijayaraghavan's paper on Borel summability [**11**].

LEMMA 5. *Under the conditions of Theorem* I, *if* $\{A_n\}$ *be a bounded sequence of numbers, and if*

$$\lim_{x\to R}\frac{1}{F(x)}\sum_{1}^{\infty}A_n g(n)e^{-G(n)}x^n = A,$$

then

$$\lim_{x\to R}\frac{1}{F_1(x)}\sum_{1}^{\infty}A_n e^{-G(n)}x^n = A.$$

Let δ be a positive number such that $0 < \delta < 1$. Let $N_1 = [(1-\delta)\xi]$ and $N_2 = [(1+\delta)\xi]$. The conditions of Lemma 1 are satisfied by these definitions of N_1 and N_2. Hence

$$\lim_{x\to R}\frac{1}{F(x)}\sum_{N_1+1}^{N_2-1}g(n)e^{-G(n)}x^n = 1, \tag{1}$$

and

$$\lim_{x\to R}\frac{1}{F(x)}\sum_{1}^{N_1}g(n)e^{-G(n)}x^n = \lim_{x\to R}\frac{1}{F(x)}\sum_{N_2}^{\infty}g(n)e^{-G(n)}x^n = 0. \tag{2}$$

Since the sequence $\{A_n\}$ is bounded, it follows that

$$\lim_{x\to R}\frac{1}{F(x)}\sum_{1}^{N_1}A_n g(n)e^{-G(n)}x^n = \lim_{x\to R}\frac{1}{F(x)}\sum_{N_2}^{\infty}A_n g(n)e^{-G(n)}x^n = 0 \tag{3}$$

and

$$\lim_{x\to R}\frac{1}{F(x)}\sum_{N_1+1}^{N_2-1}A_n g(n)e^{-G(n)}x^n = A. \tag{4}$$

As was noted under D_c we have

$$\lim_{x\to\infty}\frac{L(\lambda x)}{L(x)} = 1 \tag{5}$$

uniformly for all λ in any finite interval not containing the origin. Now if $N_1 + 1 \leqq n \leqq N_2 - 1$ then

$$(1 - \delta)\xi \leqq n \leqq (1 + \delta)\xi, \tag{6}$$

and therefore by (5)

$$1 - \epsilon(x) < \frac{L(n)}{L(\xi)} < 1 + \epsilon(x) \qquad (N_1 + 1 \leqq n \leqq N_2 - 1;\ \epsilon(x) \to 0, x \to R). \tag{7}$$

If we let $\delta_\sigma = \delta$ for $\sigma \geqq 0$ and $\delta_\sigma = -\delta$ for $\sigma < 0$, then

$$(1 - \delta_\sigma)^\sigma < \left(\frac{n}{\xi}\right)^\sigma < (1 + \delta_\sigma)^\sigma. \tag{8}$$

Combining (7) and (8) we get

$$(1 - \epsilon(x))(1 - \delta_\sigma)^\sigma < \frac{g(n)}{g(\xi)} < (1 + \epsilon(x))(1 + \delta_\sigma)^\sigma. \tag{9}$$

Since the A_n are bounded and the summability methods under consideration are regular, there is no loss of generality in assuming that the A_n are positive. Hence assuming that $A_n > 0$ we get from (9)

$$\begin{aligned}(1 - \delta_\sigma)^\sigma(1 - \epsilon(x))\frac{g(\xi)}{F(x)}\sum_{N_1+1}^{N_2-1}A_n e^{-G(n)}x^n &< \frac{1}{F(x)}\sum_{N_1+1}^{N_2-1}A_n g(n)e^{-G(n)}x^n \\ &< (1 + \delta_\sigma)^\sigma(1 + \epsilon(x))\frac{g(\xi)}{F(x)}\sum_{N_1+1}^{N_2-1}A_n e^{-G(n)}x^n.\end{aligned} \tag{10}$$

Hence

$$(1 - \delta_\sigma)^\sigma \varlimsup_{x\to R}\frac{g(\xi)}{F(x)}\sum_{N_1+1}^{N_2-1}A_n e^{-G(n)}x^n \leqq A \leqq (1 + \delta_\sigma)^\sigma \varliminf_{x\to R}\frac{g(\xi)}{F(x)}\sum_{N_1+1}^{N_2-1}A_n e^{-G(n)}x^n. \tag{11}$$

If in (1) we place $g(n) \equiv 1$, then $F(x) = F_1(x)$ and we have

$$\lim_{x\to R}\frac{1}{F_1(x)}\sum_{N_1+1}^{N_2-1}e^{-G(n)}x^n = 1. \tag{12}$$

If now we place $A_n \equiv 1$ in (11), then $A = 1$, and combining with (12) we get

$$(1 - \delta_\sigma)^\sigma \varlimsup_{x\to R}\frac{g(\xi)F_1(x)}{F(x)} \leqq 1 \leqq \varliminf_{x\to R}(1 + \delta_\sigma)^\sigma\frac{g(\xi)F_1(x)}{F(x)}. \tag{13}$$

Since δ is arbitrary, we may set $\delta_\sigma = 0$ in (13) and thus conclude that

$$\frac{g(\xi)}{F(x)} \sim \frac{1}{F_1(x)} \qquad (x \to R). \tag{14}$$

Substituting this result in (11) we have

$$\begin{aligned}(1-\delta_\sigma)^\sigma \varlimsup_{x\to R} \frac{1}{F_1(x)} \sum_{N_1+1}^{N_2-1} A_n e^{-G(n)} x^n \\ \leqq A \leqq (1+\delta_\sigma)^\sigma \varliminf_{x\to R} \frac{1}{F_1(x)} \sum_{N_1+1}^{N_2-1} A_n e^{-G(n)} x^n.\end{aligned} \tag{15}$$

If in (3) we set $g(n) \equiv 1$, then $F(x) = F_1(x)$ and we have

$$\lim_{x\to R} \frac{1}{F_1(x)} \sum_{1}^{N_1} A_n e^{-G(n)} x^n = \lim_{x\to R} \frac{1}{F_1(x)} \sum_{N_2}^{\infty} A_n e^{-G(n)} x^n. \tag{16}$$

Combining (15) and (16) we have

$$(1-\delta_\sigma)^\sigma \varlimsup_{x\to R} \frac{1}{F_1(x)} \sum_{1}^{\infty} A_n e^{-G(n)} x^n \leqq A \leqq (1+\delta_\sigma)^\sigma \varliminf_{x\to R} \frac{1}{F_1(x)} \sum_{1}^{\infty} A_n e^{-G(n)} x^n. \tag{17}$$

And since we may place $\delta_\sigma = 0$ in (17), we conclude that

$$\lim_{x\to R} \frac{1}{F_1(x)} \sum_{1}^{\infty} A_n e^{-G(n)} x^n = A. \tag{18}$$

LEMMA 6. *Let $\{A_n\}$ be a bounded sequence of numbers, $|A_n| < K$, and suppose that*

$$\lim_{x\to R} \frac{1}{F(x)} \sum_{1}^{\infty} A_n e^{-G(n)} x^n = A.$$

Let the function $A(x)$ be defined as follows:

$$\text{(i)} \qquad \begin{aligned} A(x) &= A_{[x]} + (x - [x])(A_{[x+1]} - A_{[x]}) && (x \geqq 0), \\ A(x) &= 0 && (x < 0). \end{aligned}$$

Finally, let $H(x) \sim G''(x)$ as $x \to \infty$. Then

$$\lim_{x\to\infty} \sqrt{\frac{H(x)}{2\pi}} \int_{-\infty}^{\infty} A(t+x) e^{-\frac{1}{2}t^2 H(x)}\, dt = A.$$

Since $H(x) \sim G''(x)$, we have

$$H(x) = G''(x) + \rho(x) G''(x) \qquad (\rho(x) \to 0,\ x \to \infty). \tag{1}$$

By B, we may assume that the function $\psi(x)$ of II satisfies the conditions:

$$\psi(x) = o\left\{\frac{1}{\rho(x)}\right\}, \tag{2}$$

$$\psi(x) = o\left\{\frac{1}{G''(x)}\right\} \tag{3}$$

as $x \to \infty$. Let

$$N_2 - \xi = \xi - N_1 = \left[\sqrt{\frac{\psi(\xi)}{G''(\xi)}}\right]. \tag{4}$$

From (4) and II we have

$$\sqrt{G''(N_2)} \sim \sqrt{G''(\xi)} \qquad (\xi \to \infty). \tag{5}$$

Hence

$$\lim_{\xi\to\infty} (N_2 - \xi)\sqrt{G''(N_2)} = \lim_{\xi\to\infty} (N_2 - \xi)\sqrt{G''(\xi)} = \infty. \tag{6}$$

We may then apply Lemma 5 and get

$$\lim_{x\to R} \frac{1}{F_1(x)} \sum_1^\infty A_n e^{-G(n)} x^n = A. \tag{7}$$

By (16) in Lemma 5, we have

$$\lim_{x\to R} \frac{1}{F_1(x)} \sum_{N_1+1}^{N_2-1} A_n e^{-G(n)} x^n = A. \tag{8}$$

Now

$$\frac{T_n(x)}{T_\xi(x)} = e^{-G(n)+G(\xi)+(n-\xi)\log x}, \tag{9}$$

and by Taylor's theorem

$$G(n) = G(\xi) + (n-\xi)G'(\xi) + \frac{(n-\xi)^2}{2} G''(\xi_n) \quad (\xi \leqq \xi_n \leqq n \text{ or } n \leqq \xi_n \leqq \xi). \tag{10}$$

Combining (9) and (10) we have

$$T_n = T_\xi e^{-\frac{1}{2}(n-\xi)^2 H(\xi)} \cdot e^{\frac{1}{2}(n-\xi)^2 \{H(\xi) - G''(\xi_n)\} + (n-\xi)\{\log x - G'(\xi)\}}. \tag{11}$$

Now if $N_1 \leqq n \leqq N_2$, then

$$|\xi_n - \xi| \leqq |n - \xi| \leqq N_2 - \xi = \left[\sqrt{\frac{\psi(\xi)}{G''(\xi)}}\right]. \tag{12}$$

Hence by condition II,

$$G''(\xi_n) = G''(\xi) + o\left\{\frac{G''(\xi)}{\psi(\xi)}\right\}. \tag{13}$$

Hence

$$H(\xi) - G''(\xi_n) = \rho(\xi) G''(\xi) + o\left\{\frac{G''(\xi)}{\psi(\xi)}\right\}. \tag{14}$$

Combining (14) and (2) we see that

$$H(\xi) - G''(\xi_n) = o\left\{\frac{G''(\xi)}{\psi(\xi)}\right\}. \tag{15}$$

Hence

$$(16)\qquad |(n-\xi)^2\{H(\xi)-G''(\xi_n)\}| \leqq (N_2-\xi)^2\cdot o\left\{\frac{G''(\xi)}{\psi(\xi)}\right\} = o(1)$$

uniformly for all n such that $N_1 \leqq n \leqq N_2$. Now, $\log x = G'(y)$, where $\xi = [y+1]$. Hence, since $G'(x)$ is an increasing function of x,

$$(17)\qquad 0 < G'(\xi) - \log x \leqq G'(\xi) - G'(\xi-1).$$

Hence, if $N_1 \leqq n \leqq N_2$,

$$(18)\qquad \begin{aligned} |(n-\xi)\{\log x - G'(\xi)\}| &\leqq (N_2-\xi)\{G'(\xi)-G'(\xi-1)\} \\ &\leqq (N_2-\xi)G''(\xi-1) \\ &\leqq \sqrt{\frac{\psi(\xi)}{G''(\xi)}}\cdot G''(\xi-1) \sim \sqrt{\psi(\xi)G''(\xi)}. \end{aligned}$$

Combining (18) with (3) we see that

$$(19)\qquad |(n-\xi)\{\log x - G'(\xi)\}| = o(1)$$

uniformly for all n such that $N_1 \leqq n \leqq N_2$. From (11), (16), and (19) we get

$$(20)\qquad T_n = \{1+\epsilon_n(\xi)\}T_\xi e^{-\frac{1}{2}(n-\xi)^2H(\xi)}$$

(where $\lim_{\xi\to\infty} \epsilon_n(\xi) = 0$ uniformly for all n such that $N_1 \leqq n \leqq N_2$). From (20) we conclude

$$(21)\qquad \begin{aligned} \lim_{x\to R}\frac{1}{F_1(x)}\sum_{N_1+1}^{N_2-1} A_nT_n(x) &= \lim_{x\to R}\frac{T_\xi(x)}{F_1(x)}\sum_{n=N_1+1}^{N_2-1} A_n\{1+\epsilon_n(\xi)\}e^{-\frac{1}{2}(n-\xi)^2H(\xi)} \\ &= \lim_{x\to R}\frac{T_\xi(x)}{F_1(x)}\sum_{N_1+1}^{N_2-1} A_ne^{-\frac{1}{2}(n-\xi)^2H(\xi)}. \end{aligned}$$

In particular, if we place $A_n \equiv 1$, then $A = 1$, and we get from (21)

$$(22)\qquad \lim_{x\to R}\frac{T_\xi(x)}{F_1(x)}\sum_{N_1+1}^{N_2-1} e^{-\frac{1}{2}(n-\xi)^2H(\xi)} = 1.$$

Now it is easy to show that

$$(23)\qquad \sum_{n=N_1+1}^{N_2-1} e^{-\frac{1}{2}(n-\xi)^2H(\xi)} \sim 2\int_0^\infty e^{-\frac{1}{2}t^2H(\xi)}\,dt = \sqrt{\frac{2\pi}{H(\xi)}}.$$

Thus, we see that

$$(24)\qquad \frac{T_\xi(x)}{F_1(x)} \sim \sqrt{\frac{H(\xi)}{2\pi}}.$$

And combining with (21) we have

$$(25)\qquad \lim_{\xi\to\infty}\sqrt{\frac{H(\xi)}{2\pi}}\sum_{N_1+1}^{N_2-1} A_ne^{-\frac{1}{2}(n-\xi)^2H(\xi)} = A.$$

From (25) it is possible to conclude by the sort of argument used by Hardy and Littlewood that [**1**, p. 39, Lemma 2.13]

$$\lim_{x\to\infty} \sqrt{\frac{H(x)}{2\pi}} \int_{-\infty}^{\infty} A_{[t+x]} e^{-\frac{1}{2}t^2 H(x)}\, dt = A. \tag{26}$$

And finally, by an argument which is the precise analogue of that given by Vijayaraghavan, we can conclude from (26) that [**11**, p. 322, Lemma 2]

$$\lim_{x\to\infty} \sqrt{\frac{H(x)}{2\pi}} \int_{-\infty}^{\infty} A(t+x) e^{-\frac{1}{2}t^2 H(x)}\, dt = A. \tag{27}$$

We now introduce the following lemma which was proved by Valiron [**9**, Section 11, p. 278]:

LEMMA. *Let $K(x)$ be an increasing function of x which satisfies the conditions*

$$\lim_{x\to\infty} \frac{K(x)}{\sqrt{x}} = \infty, \tag{i}$$

$$AK(x) < xK'(x) < BK(x), \tag{ii}$$

$$| x^2K''(x) | < CK(x) \tag{iii}$$

for positive A, B, C and all large x. Let $f(x)$ be a bounded function of x, and suppose that

$$\lim_{x\to\infty} \frac{1}{\sqrt{\pi x}} \int_{-\infty}^{\infty} f(t + K(x)) e^{-t^2/x}\, dt = s;$$

then

$$\lim_{x\to\infty} \sqrt{\frac{a}{\pi x}} \int_{-\infty}^{\infty} f(t + K(x)) e^{-at^2/x}\, dt = s$$

for every $a \geqq 1$.

From this lemma we obtain very easily

LEMMA 7. *If $A(x)$ is a bounded function of x, and if*

$$\lim_{x\to\infty} \sqrt{\frac{H(x)}{2\pi}} \int_{-\infty}^{\infty} A(t+x) e^{-\frac{1}{2}t^2 H(x)}\, dt = A,$$

then

$$\lim_{x\to\infty} \sqrt{\frac{aH(x)}{2\pi}} \int_{-\infty}^{\infty} A(t+x) e^{-\frac{1}{2}at^2 H(x)}\, dt = A$$

for every $a \geqq 1$, where $H(x)$ is the function defined in condition III.

To prove this, we observe that the function $K(x)$ of condition III satisfies the conditions of Valiron's lemma. Thus, if we replace the parameter x by $2/H(x)$ in this lemma, we get Lemma 7.

Vijayaraghavan has proved the following [**11**, p. 324, Lemma 4]:

LEMMA. *If δ be any positive constant and if $a \to \infty$, then*

$$I_\delta = \sqrt{\frac{a}{\pi x}} \int_{-\delta x^{\frac{1}{2}}}^{\delta x^{\frac{1}{2}}} e^{-at^2/x}\, dt \to 1$$

uniformly in x, as $a \to \infty$.

If in this lemma we make the substitution $x = 1/H(y)$, it takes the form of

LEMMA 8. *If δ is any positive constant, then*

$$I_\delta = \sqrt{\frac{aH(y)}{\pi}} \int_{-\delta(H(y))^{-\frac{1}{2}}}^{\delta(H(y))^{-\frac{1}{2}}} e^{-at^2H(y)}\, dt \to 1$$

uniformly in y, as $a \to \infty$.

Lemmas 1–8 furnish us with complete analogues of the lemmas which Vijayaraghavan uses in proving the Tauberian theorem of Borel summability. The proof of Theorem I can now be completed in the same manner as that used by Vijayaraghavan.

II

As was stated at the beginning of this paper Theorem I contains Borel summability as a special case, but not Abel summability. We turn now to a theorem which is a generalization of the Tauberian theorem of Abel summability.

THEOREM II. *Let $\Phi(x)$ be an increasing function of "regular growth", which tends to infinity with x;*

$$\Phi(x) = x^\alpha L(x) \qquad \left(\alpha \geqq 0;\ \lim_{x\to\infty} \frac{L(\lambda x)}{L(x)} = 1 \quad \textit{for every } \lambda > 0\right).$$

Let $B(x)$ be bounded in every finite interval. Let the integrals

$$\int_a^t B(u)\, d\Phi(u), \qquad \int_a^\infty e^{-t/x}\, d\left\{\int_a^t B(u)\, d\Phi(u)\right\}$$

exist as Stieltjes integrals. Then if

$$\lim_{x\to\infty} \frac{1}{\Gamma(1+\alpha)\Phi(x)} \int_a^\infty e^{-t/x}\, d\left\{\int_a^t B(u)\, d\Phi(u)\right\} = B \qquad (B \textit{ finite}),$$

and if

$$\underline{\lim}\, (B(y) - B(x)) \geqq 0, \qquad \textit{whenever } \frac{\Phi(y)}{\Phi(x)} \to 1 \quad (y \geqq x \to \infty),$$

then

$$\lim_{x\to\infty} B(x) = B.$$

LEMMA 1. *If $\Phi(N)/\Phi(x) \to 0$, then*

$$\frac{1}{\Phi(x)} \int_a^N e^{-t/x}\, d\Phi(t) \to 0.$$

For

$$(1) \qquad \frac{1}{\Phi(x)}\int_a^N e^{-t/x}\,d\Phi(t) \leqq \frac{1}{\Phi(x)}\int_a^N d\Phi(t) = \frac{\Phi(N)-\Phi(a)}{\Phi(x)} \to 0$$

by hypothesis.

LEMMA 2. *If* $\Phi(M)/\Phi(x) \to \infty$, *then*

$$\frac{1}{\Phi(x)}\int_M^\infty e^{-t/x}\,d\Phi(t) \to 0.$$

From the representation

$$(1) \qquad L(x) = c(x)e^{\int_a^x t^{-1}\epsilon(t)dt},$$

it follows that

$$(2) \qquad L(x) = o(x^\epsilon)$$

for every fixed $\epsilon > 0$, and a positive constant K exists such that if ϵ is a fixed positive number, then for all large x and $M > x$ we have

$$(3) \qquad \frac{\Phi(M)}{\Phi(x)} = \frac{c(M)}{c(x)}\left(\frac{M}{x}\right)^\alpha e^{\int_x^M t^{-1}\epsilon(t)\,dt} < K\left(\frac{M}{x}\right)^\alpha e^{\epsilon\int_x^M t^{-1}dt} = K\left(\frac{M}{x}\right)^{\alpha+\epsilon}.$$

From (3) we conclude that if $\Phi(M)/\Phi(x) \to \infty$, then

$$(4) \qquad \frac{M}{x} \to \infty.$$

From (2) we have

$$(5) \qquad \Phi(x) = x^\alpha L(x) = o(x^{\alpha+\epsilon}) \qquad (x \to \infty).$$

Integrating by parts we have

$$(6) \qquad \frac{1}{\Phi(x)}\int_M^\infty e^{-t/x}\,d\Phi(t) = \frac{\Phi(t)}{\Phi(x)}e^{-t/x}\Big]_M^\infty + \frac{1}{\Phi(x)}\int_M^\infty \frac{\Phi(t)}{x}e^{-t/x}\,dt = \frac{\Phi(M)}{\Phi(x)}e^{-M/x} + \int_M^\infty \frac{\Phi(t)}{\Phi(x)}e^{-t/x}\,d_t\left(\frac{t}{x}\right).$$

Substituting (3) in (6) we get

$$(7) \qquad \frac{1}{\Phi(x)}\int_M^\infty e^{-t/x}\,d\Phi(t) < K\left(\frac{M}{x}\right)^{\alpha+\epsilon}e^{-M/x} + \int_M^\infty \left(\frac{t}{x}\right)^{\alpha+\epsilon}e^{-t/x}\,d_t\left(\frac{t}{x}\right) = K\left(\frac{M}{x}\right)^{\alpha+\epsilon}e^{-M/x} + \int_{M/x}^\infty u^{\alpha+\epsilon}e^{-u}\,du,$$

and since, by (4), $M/x \to \infty$, we conclude from (7) that

$$\frac{1}{\Phi(x)} \int_M^\infty e^{-t/x}\, d\Phi(t) \to 0 \qquad \left(\frac{\Phi(M)}{\Phi(x)} \to \infty\right). \tag{8}$$

LEMMA 3. *If* $\Phi(N)/\Phi(x) \to 0$ *and* $\Phi(M)/\Phi(x) \to \infty$, *then*

$$\frac{1}{\Gamma(1+\alpha)\Phi(x)} \int_N^M e^{-t/x}\, d\Phi(t) \to 1.$$

It has been shown by Karamata [**4**, p. 296, Theorem I; p. 298, footnote 9] that if the function $\Phi(x)$ is a function of regular growth which tends to infinity with x, then

$$\lim_{x\to\infty} \frac{1}{\Gamma(1+\alpha)\Phi(x)} \int_a^\infty e^{-t/x}\, d\Phi(t) = 1. \tag{1}$$

If we combine (1) with Lemmas 1 and 2, Lemma 3 follows directly.

LEMMA 4. *If* $\Phi(M)/\Phi(x) \to \infty$, *then*

$$\frac{1}{\Phi(x)} \overline{\int_M^\infty} \log\left\{\frac{\Phi(t)}{\Phi(M)}\right\} e^{-t/x}\, d\Phi(t) \to 0.$$

$$\begin{aligned} \frac{1}{\Phi(x)} \overline{\int_M^\infty} \log\left\{\frac{\Phi(t)}{\Phi(M)}\right\} e^{-t/x}\, d\Phi(t) &< \frac{1}{\Phi(x)} \overline{\int_M^\infty} \frac{\Phi(t)}{\Phi(M)} e^{-t/x}\, d\Phi(t) \\ &= \frac{1}{\Phi(M)} \overline{\int_M^\infty} \frac{\Phi(t)}{\Phi(x)} e^{-t/x}\, d\Phi(t). \end{aligned} \tag{1}$$

But for all large x,

$$\frac{\Phi(t)}{\Phi(x)} < K\left(\frac{t}{x}\right)^{\alpha+\epsilon}, \tag{2}$$

and since $M/x \to \infty$, it follows that there exists a positive number δ satisfying $0 < \delta < 1$ such that, if $t \geqq M$,

$$K\left(\frac{t}{x}\right)^{\alpha+\epsilon} < e^{\delta t/x}. \tag{3}$$

Substituting (2) and (3) in (1) we get

$$\frac{1}{\Phi(x)} \overline{\int_M^\infty} \log\left\{\frac{\Phi(t)}{\Phi(M)}\right\} e^{-t/x}\, d\Phi(t) < \frac{1}{\Phi(M)} \int_M^\infty e^{-(1-\delta)t/x}\, d\Phi(t). \tag{4}$$

Integrating the right side of (4) by parts, we have

$$\begin{aligned} \frac{1}{\Phi(M)} \int_M^\infty e^{-(1-\delta)t/x}\, d\Phi(t) &= \frac{\Phi(t)}{\Phi(M)} e^{-(1-\delta)t/x} \Big]_M^\infty + \frac{1-\delta}{\Phi(M)} \int_M^\infty \frac{\Phi(t)}{x} e^{-(1-\delta)t/x}\, dt \\ &= -e^{-(1-\delta)M/x} + (1-\delta)\frac{\Phi(x)}{\Phi(M)} \int_M^\infty \frac{\Phi(t)}{\Phi(x)} e^{-(1-\delta)t/x}\, d_t\left(\frac{t}{x}\right). \end{aligned} \tag{5}$$

Since $M/x \to \infty$, and $\Phi(M)/\Phi(x) \to \infty$, it follows from (5) in precisely the same manner as in Lemma 2 that the right-hand side of (5) tends to zero. Hence we see that

$$(6) \qquad \frac{1}{\Phi(x)} \int_M^\infty \log \left\{ \frac{\Phi(t)}{\Phi(M)} \right\} e^{-t/x} \, d\Phi(t) \to 0 \qquad \left(\frac{\Phi(M)}{\Phi(x)} \to \infty \right).$$

By an argument similar to that used by Vijayaraghavan [**10**, p. 114, Lemma 6], if we make use of the fact that $\int_a^t B(u)\, d\Phi(u)$ exists as a Stieltjes integral, we can prove

LEMMA 5. *If* $\underline{\lim}\, (B(y) - B(x)) \geqq 0$ *whenever* $\Phi(y)/\Phi(x) \to 1$ $(y \geqq x \to \infty)$, *then corresponding to any positive number c there exists a positive constant* $k = k(c)$ *such that*

$$B(y) - B(x) > -\left\{ k \log \left\{ \frac{\Phi(y)}{\Phi(x)} \right\} + c \right\},$$

for all large x and $y \geqq x$.

Making use of Lemmas 1–5, which correspond to lemmas in Vijayaraghavan's paper on Abel summability [**10**], we can by an argument similar to Vijayaraghavan's prove

LEMMA 6. *Under the conditions of Theorem* II, *if*

$$\frac{1}{\Phi(x)} \int_a^\infty e^{-t/x} \, d\left\{ \int_a^t B(u)\, d\Phi(u) \right\} = O(1) \qquad (x \to \infty),$$

and if

$$\underline{\lim}\, (B(y) - B(x)) \geqq 0 \qquad \textit{whenever } \frac{\Phi(y)}{\Phi(x)} \to 1 \quad (y \geqq x \to \infty),$$

then

$$B(x) = O(1) \textit{ as } x \to \infty.$$

From Lemma 6 it follows that the function $B(x)$ of Theorem II is bounded. Hence there is no loss of generality in assuming that $B(x)$ is positive. Assuming then that $B(x)$ is positive, we write

$$(1) \qquad A(t) = \int_a^t B(u)\, d\Phi(u).$$

Since $\Phi(u)$ is an increasing function of u and $B(x)$ is positive, it follows that $A(t)$ is an increasing function of t. Thus the hypothesis of Theorem II may be put in the form

$$(2) \qquad \int_a^\infty e^{-t/x} \, d\{A(t)\} \sim \Gamma(1 + \alpha)\Phi(x) \qquad (x \to \infty),$$

where $\Phi(x) = x^\alpha L(x)$ $(\alpha \geqq 0)$, and $A(t)$ is an increasing function of t.

From (2) it follows by a theorem of Karamata that [**5**, p. 30, Theorem I]

$$(3) \qquad A(x) \sim \Phi(x) \qquad (x \to \infty),$$

i.e.,

$$\lim_{x\to\infty} \frac{1}{\Phi(x)} \int_a^x B(u)\, d\Phi(u) = B. \tag{4}$$

The following lemma now enables us to complete the proof of the theorem.

Lemma 7. *If $G(x)$ is an increasing function of x which tends to infinity with x and which satisfies the condition*

$$\lim_{x\to\infty} \frac{G(x-0)}{G(x)} = \lim_{x\to\infty} \frac{G(x+0)}{G(x)} = 1; \tag{i}$$

if

$$\lim_{x\to\infty} \frac{1}{G(x)} \int_a^x R(u)\, dG(u) = R \qquad (R \textit{ finite});$$

and if

$$\underline{\lim}\, (R(y) - R(x)) \geqq 0 \qquad \textit{whenever } \frac{G(y)}{G(x)} \to 1 \quad (y \geqq x \to \infty);$$

then

$$\lim_{x\to\infty} R(x) = R.$$

This lemma is a generalization of a lemma proved by Szász [8, p. 337, (18′)] and can be proved in the same way. There is merely one point in the proof which requires special attention. Following Szász we set up the identity

$$\begin{aligned} R(x)\{G(y) - G(x)\} &\equiv \int_x^y R(x)\, dG(t) \equiv \int_a^y R(t)\, dG(t) \\ &\quad - \int_a^x R(t)\, dG(t) - \int_x^y (R(t) - R(x))\, dG(t). \end{aligned} \tag{1}$$

To make the analogy complete, we should have to choose y so that $G(y) = (1 + \delta)G(x)$, where δ is a positive number. Since $G(x)$ is not necessarily a continuous function of x, it is not possible in general so to choose y. However, it is sufficient for the purpose of this proof if $G(y) \sim (1 + \delta)G(x)$ as $x \to \infty$. Now since $G(x)$ tends monotonically to infinity, it is clear that $y = y(x)$ can be found so that

$$G(y - 0) \leqq (1 + \delta)G(x) \leqq G(y + 0). \tag{2}$$

Combining (2) with (i) we see that $G(y) \sim (1 + \delta)G(x)$. The remainder of the proof of this lemma can be carried out in precisely the same manner as in Szász's proof.

It is easy to show that condition (i) of Lemma 7 is satisfied by the function $\Phi(x)$, since $\Phi(x)$ is of regular growth. Hence, since

$$\lim_{x\to\infty} \frac{1}{\Phi(x)} \int_a^x B(u)\, d\Phi(u) = B$$

and $\underline{\lim}\,(B(y) - B(x)) \geqq 0$ whenever $\Phi(y)/\Phi(x) \to 1$ $(y \geqq x \to \infty)$, it follows from Lemma 7 that $\lim_{x\to\infty} B(x) = B$.

References

1. G. H. Hardy and J. E. Littlewood, *Theorems concerning the summability of series by Borel's exponential method*, Palermo Rend., vol. 41 (1916), pp. 1–18.
2. J. M. Hyslop, *On the summability of series by a method of Valiron*, Proc. Edinburgh Math. Soc., (2), vol. 4 (1936), pp. 218–223.
3. J. Karamata, *Sur un mode de croissance régulière des fonctions*, Mathematica Cluj, vol. 4 (1930), pp. 38–53.
4. J. Karamata, *Neuer Beweis und Verallgemeinerung einiger Tauberian-Sätze*, Math. Zeitschr., vol. 33 (1931), pp. 294–299.
5. J. Karamata, *Neuer Beweis und Verallgemeinerung der Tauberschen Sätze, welche die Laplacesche und Stieltjessche Transformation betreffen*, Journ. für Math., vol. 164 (1931), pp. 27–39.
6. R. Schmidt, *Über divergente Folgen und lineare Mittelbildungen*, Math. Zeitschr., vol. 22 (1925), pp. 89–152.
7. R. Schmidt, *Die Umkehrsätze des Borelschen Summierungsverfahrens*, Schriften der König. Gel. Gesell., vol. 1 (1925), pp. 205–256.
8. O. Szász, *Verallgemeinerung und neuer Beweis einiger Sätze Tauberscher Art*, Sitzungsb. d. math.-phys. Klasse d. Akad. d. Wiss., München (1929), pp. 325–340.
9. G. Valiron, *Remarques sur la sommation des séries divergentes par les méthodes de M. Borel*, Palermo Rend., vol. 42 (1917), pp. 267–284.
10. T. Vijayaraghavan, *A Tauberian theorem*, Journ. Lond. Math. Soc., vol. 1 (1926), pp. 113–120.
11. T. Vijayaraghavan, *A theorem concerning the summability of series by Borel's method*, Proc. Lond. Math. Soc., vol. 27 (1927), pp. 316–326.

Brown University.

ASYMPTOTIC EXPRESSIONS FOR THE ZEROS OF GENERALIZED LAGUERRE POLYNOMIALS AND WEBER FUNCTIONS

By Vivian Eberle Spencer

Introduction. It is the purpose of this paper to apply the close relationship between Hermite and Laguerre polynomials to find asymptotic expressions for the zeros $\{{}_Lx_{in}\}$ of the generalized Laguerre polynomial $L_n(x, \alpha)$. The results obtained for the largest zero ${}_Lx_{nn}$ are, we believe, new; for the other zeros the expressions are essentially equivalent to those obtained by Winston;[1] the method of procedure, however, in every case is new and fruitful. Hermite functions $h(x, n)$ are the special case of Weber's parabolic cylinder functions $w(x, n)$ obtained when the boundary conditions $w(\pm\infty, n) = O(x^n e^{-\frac{1}{2}x^2})$ are imposed. The argument is simplified and the order of some of the results improved by the introduction of the latter functions for non-integral n. Hence we are led to a discussion of the zeros $\{{}_wx_{in}\}$ of $w(x, n)$. By an elementary application of Sturm's theory Milne's[2] properties of the zeros $\{{}_wx_{in}\}$ of the standard solution are obtained, and similar properties for the zeros of the solution converging to zero as x approaches minus infinity are developed. The application of the known asymptotic expressions for the zeros of Hermite polynomials is shown to give directly bounds and asymptotic expressions for the zeros of Weber functions for n an arbitrary positive number. A sequence of Weber functions $\omega(x, \alpha, n)$ is associated with every sequence of Laguerre functions $l(x, \alpha, n)$, and definite separation and asymptotic relations are obtained between the zeros of $\omega(x, \alpha, n)$ and $l(x, \alpha, n)$. As a consequence of these relations asymptotic expressions for all zeros of $L_n(x, \alpha)$, for any $\alpha > 0$, are obtainable immediately from any asymptotic expression for the zeros of the Hermite polynomial $H_n(x)$, or for the zeros of the Laguerre polynomial proper $L_n(x, 1)$. Neumann's[3] bounds for the zeros of $L_n(x, 1)$ and Zernike's[4] asymptotic expression for the largest zero of $H_n(x)$ are then applied to $L_n(x, \alpha)$.

1. **An asymptotic expression for ${}_Lx_{nn}$ for $\frac{1}{2} \leqq \alpha \leqq \frac{3}{2}$.** We consider the Hermite and Laguerre polynomials satisfying respectively the differential equations

$$H_n''(x) - 2xH_n'(x) + 2nH_n(x) = 0, \tag{1}$$

$$xL_n''(x, \alpha) + (\alpha - x)L_n'(x, \alpha) + nL_n(x, \alpha) = 0 \qquad (\alpha > 0). \tag{2}$$

Received January 4, 1937; in revised form, July 8, 1937.

[1] C. Winston, *On mechanical quadratures formulae involving the classical orthogonal polynomials*, Annals of Math., vol. 35 (1934), pp. 658–677.

[2] A. Milne, *On the roots of the confluent hypergeometric functions*, Proc. Edinburgh Math. Soc., vol. 33 (1915), pp. 48–64.

[3] E. R. Neumann, *Beiträge zur Kenntnis der Laguerreschen Polynome*, Jahresbericht der Deutschen Math.-Vereinigung, vol. 30 (1921), pp. 15–35.

[4] F. Zernike, *Eine asymptotische Entwicklung für die grösste Nullstelle der Hermiteschen Polynome*, Amsterdam Academy, Proc. of Sec. Sc., vol. 34 (1931), pp. 673–680.

The polynomials $H_n(x)$ and $L_n(x, \alpha)$ satisfy the relations

$$H_{2n}(x) \equiv L_n(x^2, \tfrac{1}{2}); \qquad H_{2n+1}(x) \equiv xL_n(x^2, \tfrac{3}{2}). \tag{3}$$

Zernike's asymptotic expression for the largest zero ${}_Hx_{nn}$ of $H_n(x)$ is[5]

$$\begin{aligned} {}_Hx_{nn} &= (2n+1)^{1/2} - 1.8557571\,(2n+1)^{-1/6} - 0.3443834\,(2n+1)^{-5/6} \\ &\quad - 0.168715\,(2n+1)^{-3/2} - 0.151965\,(2n+1)^{-13/6} + O\{(2n+1)^{-17/6}\}. \end{aligned} \tag{4}$$

Applying (3), we obtain immediately an asymptotic expression for the largest zero ${}_Lx_{nn}$ of $L_n(x, \alpha)$ for $\alpha = \frac{1}{2}, \frac{3}{2}$:

$$\begin{cases} (\alpha = \tfrac{1}{2}) \quad {}_Lx_{nn} = 4n + 1 - 3.7115142\,(4n+1)^{1/3} + 2.7550676\,(4n+1)^{-1/3} \\ \qquad + 0.940754\,(4n+1)^{-1} + 0.440858\,(4n+1)^{-5/3} \\ \qquad + O\{(4n+1)^{-7/3}\} \equiv A(n, \tfrac{1}{2}), \\ (\alpha = \tfrac{3}{2}) \quad {}_Lx_{nn} = 4n + 3 - 3.7115142\,(4n+3)^{1/3} + 2.7550676\,(4n+3)^{-1/3} \\ \qquad + 0.940754\,(4n+3)^{-1} + 0.440858\,(4n+3)^{-5/3} \\ \qquad + O\{(4n+3)^{-7/3}\} \equiv A(n, \tfrac{3}{2}). \end{cases} \tag{5}$$

Markoff's[6] theorem, applied to Laguerre polynomials, gives

$$\frac{\partial\, {}_Lx_{in}}{\partial \alpha} > 0 \qquad (i = 1, 2, \cdots, n). \tag{6}$$

Then (5) and (6) yield

$$\begin{cases} (\alpha > \tfrac{1}{2}) & {}_Lx_{nn} > A(n, \tfrac{1}{2}), \\ (0 < \alpha < \tfrac{3}{2}) & {}_Lx_{nn} < A(n, \tfrac{3}{2}). \end{cases} \tag{7}$$

Hence

$$(\tfrac{1}{2} < \alpha < \tfrac{3}{2}) \quad {}_Lx_{nn} = 4n + 2\alpha - 3.7115142\,(4n + 2\alpha)^{1/3} + O(1) \qquad (O(1) < 2). \tag{8}$$

Note that (8) includes the case of Laguerre polynomials proper, corresponding to $\alpha = 1$.[7]

A more delicate analysis is necessary to extend these results to any $\alpha > 0$.

[5] The remainder term does not appear explicitly in Zernike's expression, but is implied in his argument since he proves that an asymptotic expansion for ${}_Hx_{nn}$ in powers of $\gamma = 2n + 1$ exists.

[6] J. Shohat, *Théorie générale des polynomes orthogonaux de Tchebichef*, Mémorial des Sciences Math., Fasc. 66, p. 39.

[7] By a method based on Zernike's argument for Hermite polynomials, but not utilizing the intimate relation between h and l apparent in (10) and (21) below, W. Hahn, *Die Nullstellen der Laguerreschen und Hermiteschen Polynome*, Dissertation, Berlin, 1933, was able to show that $4n + 2\alpha - c_1(4n + 2\alpha)^{\frac{1}{3}} < {}_Lx_{nn} < 4n + 2\alpha - c_2(4n + 2\alpha)^{\frac{1}{3}}$ ($\alpha > 0$; $c_1, c_2 > 0$), for n sufficiently large. (8), which has been obtained from Markoff's theorem with scarcely any argument, is, for $\frac{1}{2} < \alpha < \frac{3}{2}$, a better result.

2. **The zeros of Weber's parabolic cylinder functions.** The Weber equation may be written in the form[8]

$$w'' + (2n + 1 - x^2)w = 0. \tag{9}$$

When the boundary conditions $w(\pm\infty, n) = O(x^n e^{-\frac{1}{2}x^2})$ are imposed, the only solutions of (9) are the Hermite functions $\{h(x, \nu)\}$, $h(x, \nu) \equiv H_\nu(x)e^{-\frac{1}{2}x^2}$, satisfying the differential equation

$$h'' + (2\nu + 1 - x^2)h = 0 \qquad (\nu = 0, 1, 2, \cdots). \tag{10}$$

To show this, assume that there exists for $n \neq \nu$ a solution of (9) satisfying these boundary conditions. Multiply (9) by h, (10) by w, subtract and integrate between any two limits a and b

$$(w'h - wh')\Big|_a^b = 2(\nu - n)\int_a^b wh\,dx. \tag{11}$$

Then letting $(a, b) \equiv (-\infty, \infty)$, we have $\int_{-\infty}^{\infty} wh\,dx = 0$, a result which contradicts the completeness of the set $\{h(x, \nu)\}$.

Write (9) and (10) in the form

$$w'' - \sigma(x, n)w = 0, \qquad \sigma(x, n) \equiv \sigma(n) = x^2 - 2n - 1, \tag{12}$$

$$h'' - \rho(x, \nu)h = 0, \qquad \rho(x, \nu) \equiv \rho(\nu) = x^2 - 2\nu - 1. \tag{13}$$

Then $\rho(\nu - 1) > \sigma(n) > \rho(\nu)$ for $\nu - 1 < n < \nu$. Hence, as a consequence of Sturm's fundamental theorem, the zeros $\{{}_wx_{in}\}$ of $w(x, n)$ separate the zeros $\{{}_Hx_{i\nu-1}\}$ of $h(x, \nu - 1)$ and are separated by the zeros $\{{}_Hx_{i\nu}\}$ of $h(x, \nu)$. It follows that $w(x, n)$ has at least $\nu - 2$ and at most $\nu + 1$ zeros.

Consider (11) with $h(x, \nu)$ replaced by $h(x, \nu - 1)$, and $(a, b) \equiv ({}_Hx_{\nu-1\nu-1}, k) \equiv (\tau, k)$, where k is an arbitrary constant to be properly chosen. We have

$$w'(k)h(k) - w(k)h'(k) + w(\tau)h'(\tau) = 2(\nu - 1 - n)\int_\tau^k wh\,dx. \tag{14}$$

If $w'(x)h(x)$ and $w(x)h'(x) \to 0$ as $x \to \infty$, choose $k = \infty$. Since $h(x) > 0$ for $x > \tau$, $h'(\tau) > 0$, then $n > \nu - 1$ implies that for both members of (14) to have the same sign it is necessary that ${}_Hx_{\nu-1\nu-1} < {}_wx_{mn}$, where ${}_wx_{mn}$ is the largest zero of $w(x, n)$.

But if either $w'(x)h(x)$ or $w(x)h'(x)$ does not approach zero as $x \to \infty$, then since $h(x)$ and $h'(x) \to 0$ as $x \to \infty$, $w(x)w'(x) \to \infty$ as $x \to \infty$ and hence for some value k of x sufficiently large $h'(k) < 0$ and $w'(k)w(k) > 0$. Choosing this value for k in (14), and supposing sgn $w(x)$ constant for $x > \tau$, then for $n > \nu - 1$ the sign of the left member is determined as sgn $w(x)$ and the sign of the right member as $-$sgn $w(x)$. This leads to a contradiction. Hence whatever the asymptotic behavior of $w(x)$ we have ${}_Hx_{\nu-1\nu-1} < {}_wx_{mn}$.

[8] This is obtainable from the standard form $D_n'' + (n + \frac{1}{2} - \frac{1}{4}x^2)D_n = 0$ by the substitution $w(x, n) = D_n(2^{\frac{1}{2}}x)$.

A similar argument applied to the interval $(a, b) \equiv (k, {}_Hx_{1\nu-1})$ leads to the inequality ${}_wx_{1n} < {}_Hx_{1\nu-1}$, whatever the asymptotic behavior of $w(x)$. Hence $w(x, n)$ has either ν or $\nu + 1$ zeros.

Consider now the standard solution of the Weber equation, defined by Whittaker[9] and discussed by Milne,[10] for which $w(x) = O(e^{-\frac{1}{4}x^2}x^p)$, p a constant, x large and positive. In (11) let $(a, b) \equiv ({}_wx_{mn}, \infty) \equiv (\tau_1, \infty)$.

$$-w'(\tau_1)h(\tau_1) = 2(\nu - n)\int_{\tau_1}^{\infty} wh\,dx. \tag{15}$$

Since $w'(\tau_1)w(x) > 0$ for $x > {}_wx_{mn}$, then $\nu > n$ implies that for both members of (15) to have the same sign it is necessary that ${}_wx_{mn} < {}_Hx_{\nu\nu}$, and hence that $w(x, n)$ has exactly ν zeros. Moreover, since the zeros of $w(x, n)$ are separated by the zeros of $h(x, \nu)$, these results imply ${}_wx_{1n} < {}_Hx_{1\nu}$.

For the solution of the Weber equation for which $w(x)$ and $w'(x) \to 0$ as $x \to -\infty$ a similar argument leads to the inequality ${}_Hx_{1\nu} < {}_wx_{mn}$, and hence to the conclusion that this solution has exactly ν zeros, and that ${}_Hx_{\nu\nu} < {}_wx_{mn}$.[11]

Since the conclusions of the two preceding paragraphs cannot hold simultaneously, it is clear that a solution of (9) for non-integral values of n cannot converge to zero at both $+\infty$ and $-\infty$.

Summing up our results, and recalling again Sturm's oscillation theorem, we have:

1. Milne's properties of the zeros $\{{}_wx_{in}\}$ of the *standard solution* of the Weber equation are: *If ν is an integer and $\nu - 1 < n \leqq \nu$, then $w(x, n)$ has ν zeros. If ν is even, $\frac{1}{2}\nu$ zeros are positive and $\frac{1}{2}\nu$ negative. If ν is odd, $\frac{1}{2}(\nu - 1)$ zeros are positive and $\frac{1}{2}(\nu + 1)$ negative or zero (zero only for $n = \nu$). Moreover, as n increases all zeros of $w(x, n)$ increase.*

2. For the *solution for which $w(x)$ and $w'(x) \to 0$ as $x \to -\infty$: If ν is an integer and $\nu - 1 < n \leqq \nu$, then $w(x, n)$ has ν zeros. If ν is even, $\frac{1}{2}\nu$ zeros are positive and $\frac{1}{2}\nu$ negative. If ν is odd, $\frac{1}{2}(\nu + 1)$ zeros are positive and $\frac{1}{2}(\nu - 1)$ negative or zero (zero only for $n = \nu$). Moreover, as n increases all zeros of $w(x, n)$ decrease.*

3. For any solution of the Weber equation: *If ν is an integer and $\nu - 1 < n \leqq \nu$, then $w(x, n)$ has either ν or $\nu + 1$ zeros. If ν is even, at least $\frac{1}{2}\nu$ zeros are positive and $\frac{1}{2}\nu$ negative. If ν is odd, at least $\frac{1}{2}(\nu - 1)$ zeros are positive and $\frac{1}{2}(\nu - 1)$ negative.*

These properties of the zeros of Weber functions enable us to extend at once bounds for the zeros of Hermite polynomials to the zeros of any Weber function. In particular, for the *standard solution* of (9), defining ${}_Hx_{0\nu-1}$ as $-\infty$,

$$ {}_Hx_{i-1\nu-1} < {}_wx_{in} < {}_Hx_{i\nu} \qquad (\nu - 1 < n < \nu;\ i = 1, 2, \cdots, \nu). \tag{16}$$

[9] E. T. Whittaker, *On the functions associated with the parabolic cylinder in harmonic analysis*, Proc. London Math. Soc., vol. 35 (1903), pp. 417–427.

[10] A. Milne, loc. cit.

[11] Incidentally, this solution is obtained from the standard solution by changing x into $-x$.

Thus, Winston's bounds for the zeros of Hermite polynomials give for $\nu - 1 < n < \nu$

$$
(17)\quad \left.\begin{aligned}
&\frac{2i-\nu-2+2\cdot 6^{\frac12}}{2^{\frac12}\nu^{\frac12}} < {}_w x_{in} < \frac{2^{\frac12}(2i-\nu+1)}{(\nu+1)^{\frac12}} \qquad (i = \tfrac12\nu+2, \cdots, \nu)\\
&\frac{\nu-2i+2^{\frac12}}{2^{\frac12}(\nu+2)^{\frac12}} < |{}_w x_{in}| < \frac{2^{\frac12}(\nu-2i+7)}{(\nu+1)^{\frac12}} \qquad (i = 2, \cdots, \tfrac12\nu)\\
&{}_w x_{1n} < -\frac{\nu-2+2^{\frac12}}{2^{\frac12}(\nu+2)^{\frac12}}; \qquad 0 < {}_w x_{\frac12\nu+1, n} < \frac{3\cdot 2^{\frac12}}{(\nu+1)^{\frac12}}
\end{aligned}\right\}\ \nu \text{ even}
$$

$$
\left.\begin{aligned}
&\frac{2i-\nu-3+2^{\frac12}}{2^{\frac12}(\nu+1)^{\frac12}} < {}_w x_{in} < \frac{2^{\frac12}(2i-\nu+4)}{(\nu+2)^{\frac12}} \qquad (i = \tfrac12(\nu+3), \cdots, \nu)\\
&\frac{\nu-2i+1+2\cdot 6^{\frac12}}{2^{\frac12}(\nu+1)^{\frac12}} < |{}_w x_{in}| < \frac{2^{\frac12}(\nu-2i+4)}{\nu^{\frac12}} \qquad (i = 2, \cdots, \tfrac12(\nu-1))\\
&{}_w x_{1n} < -\frac{\nu-1+2\cdot 6^{\frac12}}{2^{\frac12}(\nu+1)^{\frac12}}; \qquad -\frac{3\cdot 2^{\frac12}}{\nu^{\frac12}} < {}_w x_{\frac12(\nu+1)n} < 0
\end{aligned}\right\}\ \nu \text{ odd}
$$

Moreover, the *standard solution* of the Weber equation, like the Hermite functions, vanishes together with its derivatives for $x \to \infty$, and their differential equations are identical in form. These being the only properties of Hermite functions employed by Zernike in obtaining his asymptotic expression for ${}_H x_{\nu\nu}$, we may at once write

$$
(18)\quad {}_w x_{\nu n} = (2n+1)^{1/2} - 1.8557571\,(2n+1)^{-1/6} - 0.3443834\,(2n+1)^{-5/6} - 0.168715\,(2n+1)^{-3/2} - 0.151965\,(2n+1)^{-13/6} + O\{(2n-1)^{-17/6}\}.
$$

3. **Laguerre functions and associated Weber functions.** In (2) let $l(x, \alpha, n) \equiv l \equiv x^{\alpha-\frac12} e^{-\frac12 x^2} L_n(x^2, \alpha)$, a transformation carrying reals into reals for any positive α and x. Then,

$$
\begin{aligned}
&(19)\quad l'' + \left(4n + 2\alpha - x^2 - \frac{(\alpha-\frac12)(\alpha-\frac32)}{x^2}\right) l = 0,\\
&(19\text{a})\quad l'' + (4n + 2\alpha - x^2) l = \frac{(\alpha-\frac12)(\alpha-\frac32)}{x^2}\, l,
\end{aligned}
\qquad \left(\begin{gathered}\alpha > 0;\\ l(\pm\infty) = O(x^{2n+\alpha-\frac12}e^{-\frac12 x^2})\end{gathered}\right).
$$

We are thus led to consider the equation

$$
(20)\qquad \omega'' + (4n + 2\alpha - x^2)\omega = 0 \qquad (\alpha > 0),
$$

upon which we wish to impose a boundary condition of the form $\omega(+\infty) = O(x^k e^{-\frac12 x^2})$, $k > 0$. But (20) is the Weber equation with $2n + 1$ replaced by $4n + 2\alpha$; and the boundary conditions are those of the standard solution. Hence, a solution of (20), satisfying these boundary conditions, exists for every n and α; and for n and α such that

$$
(21)\qquad \nu - 1 < \frac{4n + 2\alpha - 1}{2} \leqq \nu, \qquad \nu \text{ a positive integer,}
$$

$\omega(x, \alpha, n)$ has exactly ν zeros. The set of functions $\{\omega(x, \alpha, n)\}$ so defined for given n and α shall be known as the *associated Weber functions* corresponding to the Laguerre functions $\{l(x, \alpha, n)\}$.

Condition (21) may be rewritten as $\frac{1}{4}(2\nu - 2\alpha - 1) < n < \frac{1}{4}(2\nu - 2\alpha + 1)$. The characteristic numbers giving rise to the Laguerre functions are $n = 0, \frac{1}{2}, 1, \frac{3}{2}, \cdots$. The corresponding Laguerre functions have respectively $0, 1, 2, 3, \cdots$ zeros, arranged symmetrically with respect to the origin. Let us consider the associated Weber functions $\omega(x, \alpha, n)$, $n = 0, \frac{1}{2}, 1, \frac{3}{2}, \cdots$. If $N(x)$ denote the next integer $\geqq x$, these Weber functions have respectively $N\left(\frac{2\alpha - 1}{2}\right), N\left(\frac{2\alpha + 1}{2}\right), N\left(\frac{2\alpha + 3}{2}\right), \cdots$ zeros. Hence, the number of zeros of $l(x, \alpha, n)$ equals the number of zeros of $\omega(x, \alpha, n)$ for $0 < \alpha \leqq \frac{1}{2}$, and in general for any $\alpha > 0$ the number of zeros of $\omega(x, \alpha, n)$ exceeds the number of zeros of $l(x, \alpha, n)$ by $N\left(\frac{2\alpha - 1}{2}\right)$.

Write (19) and (20) in the form

$$l'' - \rho(x, \alpha, n)l = 0, \qquad \rho(x, \alpha, n) \equiv \rho \equiv x^2 - 4n - 2\alpha + \frac{(\alpha - \frac{1}{2})(\alpha - \frac{3}{2})}{x^2}, \tag{22}$$

$$\omega'' - \sigma(x, \alpha, n)\omega = 0, \qquad \sigma(x, \alpha, n) \equiv \sigma \equiv x^2 - 4n - 2\alpha. \tag{23}$$

Then for every x and n

$$\begin{aligned} &\rho \leqq \sigma && (\tfrac{1}{2} \leqq \alpha \leqq \tfrac{3}{2}), \\ &\rho > \sigma && (0 < \alpha < \tfrac{1}{2}, \text{ or } \alpha > \tfrac{3}{2}). \end{aligned} \tag{24}$$

Hence Sturm's fundamental theorem gives: l oscillates more rapidly than ω for $\frac{1}{2} < \alpha < \frac{3}{2}$, and ω oscillates more rapidly than l for $0 < \alpha < \frac{1}{2}$, or $\alpha > \frac{3}{2}$ ($\omega \equiv l$ for $\alpha = \frac{1}{2}, \frac{3}{2}$).

Denote the zeros of $l(x, \alpha, n)$ by $\{\zeta_i\}$, $i = 1, 2, \cdots, n$, and the zeros of $\omega(x, \alpha, n)$ by $\{\delta_i\}$, $i = 1, 2, \cdots, m$, where the zeros in each set are arranged in increasing order of magnitude.

Multiply (19a) by ω, (20) by l, subtract and integrate,

$$(l'\omega - l\omega')\Big|_k^\infty = (\alpha - \tfrac{1}{2})(\alpha - \tfrac{3}{2}) \int_k^\infty \frac{l\omega}{x^2}\, dx, \tag{25}$$

where k is an arbitrary constant to be properly chosen. Since $\omega, \omega', l, l' \to 0$ as $x \to \infty$ (25) becomes

$$l'(k)\omega(k) - l(k)\omega'(k) = -(\alpha - \tfrac{1}{2})(\alpha - \tfrac{3}{2}) \int_k^\infty \frac{l\omega}{x^2}\, dx. \tag{26}$$

First, let $k = \zeta_n$; (26) reduces to

$$l'(\zeta_n)\omega(\zeta_n) = -(\alpha - \tfrac{1}{2})(\alpha - \tfrac{3}{2}) \int_{\zeta_n}^\infty \frac{l\omega}{x^2}\, dx. \tag{27}$$

ζ_n is not a multiple zero of $l(x)$ and $l(x)$ remains positive for $x > \zeta_n$; hence $l'(\zeta_n) > 0$. This for $0 < \alpha < \frac{1}{2}$ or $\alpha > \frac{3}{2}$ implies $\zeta_n < \delta_m$. Next, let $k = \delta_m$, by similar reasoning

$$l(\delta_m)\omega'(\delta_m) = (\alpha - \tfrac{1}{2})(\alpha - \tfrac{3}{2}) \int_{\delta_m}^{\infty} \frac{l\omega}{x^2}\, dx. \tag{28}$$

But $\omega(x)$ has no multiple zeros and $\omega'(\delta_m)\omega(x) > 0$ for $x > \delta_m$. For $\frac{1}{2} < \alpha < \frac{3}{2}$ this implies $\delta_m < \zeta_n$. Hence, we have

THEOREM I. *The zeros $\{\zeta_i\}$ of the Laguerre functions $l(x, \alpha, n)$ and the zeros $\{\delta_i\}$ of the associated Weber functions $\omega(x, \alpha, n)$ satisfy the following separation relations*

$$\left\{\begin{array}{ll} (0 < \alpha < \frac{1}{2}) & \delta_{i-1} < \zeta_i < \delta_i, \\ (\frac{1}{2} < \alpha < \frac{3}{2}) & \zeta_{i-1} < \delta_{i+1} < \zeta_i, \\ (\alpha > \frac{3}{2}) & \delta_{i-1} < \zeta_i < \delta_{i+N\left(\frac{2\alpha-1}{2}\right)} \leqq \delta_m, \end{array}\right. \qquad \left\{\begin{array}{ll} i = \dfrac{n}{2} + 1, \cdots, n, & n \text{ even}, \\ i = \dfrac{n+3}{2}, \cdots, n, & n \text{ odd}. \end{array}\right. \tag{29}$$

4. **Asymptotic expressions for the zeros of $L_n(x, \alpha)$.** The results of §2 combined with Theorem I give us a method whereby results established for the zeros of Hermite polynomials immediately yield results for the zeros of generalized Laguerre polynomials for any $\alpha > 0$. The zeros of Hermite functions are identical with the zeros $\{{}_Hx_{in}\}$ of the corresponding Hermite polynomials. Between any two of these zeros there lies one and only one zero of the set $\{{}_wx_{im}\}$ of zeros of the standard Weber function $w(x, m)$ for which $n = N(m)$. But the associated Weber functions, $\omega(x, \alpha, n)$, a particular case of $w(x, m)$, lead by Theorem I at once to relations for the zeros $\{\zeta_i\}$ of the Laguerre functions $l(x, \alpha, n)$. Results for the zeros $\{{}_Lx_{in}\}$ of $L_n(x, \alpha)$ are then obtained by squaring the $\{\zeta_i\}$ and replacing n by $\frac{1}{2}n$. Moreover, our method permits us to extend results for Laguerre polynomials proper, $\alpha = 1$, to the generalized Laguerre polynomials. In fact Winston,[12] applying Markoff's theorem, has indicated the following method for reducing results for Laguerre polynomials ($\alpha = 1$) to results for Hermite polynomials.

To indicate the dependence of ${}_Lx_{in}$ on α write ${}_Lx_{in} = {}_Lx_{in}(\alpha)$. Then relations (3) and (6) give

$$\begin{array}{rl} -{}_Hx_{2n+1-i,2n} \equiv {}_Hx_{i,2n} \equiv \sqrt{{}_Lx_{i-n,n}(\frac{1}{2})} < \sqrt{{}_Lx_{i-n,n}(1)} & \\ & (i = n+1, n+2, \cdots, 2n) \\ -{}_Hx_{2n+2-i,2n+1} \equiv {}_Hx_{i,2n+1} \equiv \sqrt{{}_Lx_{i-n-1,n}(\frac{3}{2})} > \sqrt{{}_Lx_{i-n-1,n}(1)} & \\ & (i = n+2, n+3, \cdots, 2n+1). \end{array} \tag{30}$$

[12] C. Winston, loc. cit., p. 676.

Hence, since ${}_Hx_{i,2n} > {}_Hx_{i,2n+1}$ $(1 < i < 2n)$,

$$\sqrt{{}_Lx_{i-n-1,n}(1)} < -{}_Hx_{2n+1-i,2n} \equiv {}_Hx_{i,2n} < \sqrt{{}_Lx_{i-n,n}(1)} \tag{31}$$
$$(i = n+1, n+2, \cdots, 2n),$$

where we define the number ${}_Lx_{0,n}(1) = 0$.

Thus, extending Neumann's[13] results for Laguerre polynomials $\alpha = 1$ to generalized Laguerre polynomials the process of the preceding paragraph gives

$$\frac{i-n}{2(n+1)^{\frac{1}{2}}} < -{}_Hx_{2n+1-i,2n} \equiv {}_Hx_{i,2n} < \frac{2(i-n+1)}{(n+1)^{\frac{1}{2}}} \tag{32}$$
$$(i = n+1, n+2, \cdots, 2n).$$

If we apply §2 and Theorem I, (32) gives[14]

$$\frac{(i-2)^2}{4(n+1)} < {}_Lx_{in} < \frac{4(i+1)^2}{n+1} \qquad (0 < \alpha < \tfrac{1}{2}, i = 2, 3, \cdots, n),$$
$$\frac{(i-1)^2}{4(n+1)} < {}_Lx_{in} < \frac{4(i+2)^2}{n+1} \qquad (\tfrac{1}{2} < \alpha < \tfrac{3}{2}, i = 1, 2, \cdots, n-1), \tag{33}$$
$$\frac{(i-2)^2}{4(n+1)} < {}_Lx_{in} < \frac{4\left(i + N\left(\frac{2\alpha+1}{2}\right)\right)^2}{n+1} \qquad (\alpha > \tfrac{3}{2}, i = 2, 3, \cdots, n-1).$$

Replacing $2n + 1$ by $4n + 2\alpha$ in the asymptotic expression (18) for ${}_wx_{nn}$, we obtain as an asymptotic expression for the largest zero δ_m of $\omega(x, \alpha, n)$

$$\begin{aligned} \delta_m = {} & (4n+2\alpha)^{1/2} - 1.8557571\,(4n+2\alpha)^{-1/6} - 0.3443834\,(4n+2\alpha)^{-5/6} \\ & - 0.168715\,(4n+2\alpha)^{-3/2} - 0.151965\,(4n+2\alpha)^{-13/6} \\ & + O\{(4n+2\alpha)^{-17/6}\} \equiv a(\alpha, n). \end{aligned} \tag{34}$$

Let us now investigate the proximity, asymptotically, of δ_m to ζ_m. Replace n in (20) by $n' < n$ and ω by ω_1,

$$\omega_1'' + (4n' + 2\alpha - x^2)\omega_1 = 0. \tag{35}$$

Multiply (19a) by ω_1, (35) by l, subtract, and integrate,

$$(l'\omega_1 - l\omega_1')\Big|_k^\infty + 4(n-n')\int_k^\infty l\omega_1\,dx = (\alpha - \tfrac{1}{2})(\alpha - \tfrac{3}{2})\int_k^\infty \frac{l\omega}{x^2}\,dx, \tag{36}$$

whence

$$l(k)\omega_1'(k) - l'(k)\omega_1(k) = -4(n-n')\int_k^\infty l\omega_1\,dx + (\alpha - \tfrac{1}{2})(\alpha - \tfrac{3}{2})\int_k^\infty \frac{l\omega_1}{x^2}\,dx.$$

[13] E. R. Neumann, *Beiträge zur Kenntnis der Laguerreschen Polynome*, Jahresbericht der Deutschen Math. Vereinigung, vol. 30 (1921), pp. 15–35.

[14] C. Winston, loc. cit., p. 675, has obtained results essentially equivalent to these by a geometric argument similar to Neumann's.

Proceeding as in the previous section, let $k = \delta_m$, the largest zero of ω_1.

$$(37)\qquad l(\delta_{m'})\omega_1'(\delta_{m'}) = -4(n-n')\int_{\delta_{m'}}^{\infty} l\omega_1\,dx + (\alpha-\tfrac{1}{2})(\alpha-\tfrac{3}{2})\int_{\delta_{m'}}^{\infty}\frac{l\omega_1}{x^2}\,dx.$$

Consider (37) for $0 < \alpha < \frac{1}{2}$ or $\alpha > \frac{3}{2}$. Assume $l(x) > 0$ for $x > \delta_{m'}$. Since $\omega_1(\delta_{m'})\omega_1'(x) > 0$ for $x > \delta_{m'}$, (37) implies

$$4(n-n')\int_{\delta_{m'}}^{\infty} l\omega_1\,dx < (\alpha-\tfrac{1}{2})(\alpha-\tfrac{3}{2})\int_{\delta_{m'}}^{\infty}\frac{l\omega_1}{x^2}\,dx$$

$$< \frac{(\alpha-\frac{1}{2})(\alpha-\frac{3}{2})}{\delta_{m'}^2}\int_{\delta_{m'}}^{\infty} l\omega_1\,dx,$$

$$4(n-n') < \frac{(\alpha-\frac{1}{2})(\alpha-\frac{3}{2})}{\delta_{m'}^2}.$$

This inequality will be contradicted if

$$(38)\qquad n' < n - \frac{(\alpha-\frac{1}{2})(\alpha-\frac{3}{2})}{4\delta_{m'}^2}.$$

Hence, for n' satisfying (38), $\delta_{m'} < \zeta_n$. Combining this result with (34), we have

$$(39)\qquad a(\alpha, n') \leqq \zeta_n \leqq a(\alpha, n) \qquad (0 < \alpha < \tfrac{1}{2} \text{ or } \alpha > \tfrac{3}{2}).$$

Similarly, replace n in (19a) by $n' < n$. Then reasoning analogous to that used in deriving (39) gives

$$(40)\qquad \zeta_{n'} < \delta_m \text{ for } n' < n - \frac{|(\alpha-\frac{1}{2})(\alpha-\frac{3}{2})|}{4\delta_{m'}^2} \qquad (\tfrac{1}{2} < \alpha < \tfrac{3}{2}).$$

Hence, for n' satisfying (40)

$$(41)\qquad a(\alpha, n') \leqq \zeta_{n'} \leqq a(\alpha, n) \qquad (\tfrac{1}{2} < \alpha < \tfrac{3}{2}).$$

But the first term of (34) which will be affected by replacing n by $n + O(n^{-1})$ is the term in $n^{-3/2}$, hence (33), (34), (39) and (41) lead to

THEOREM II. *If $\{{}_Lx_{in}\}$ denote the zeros of the generalized Laguerre polynomial $L_n(x, \alpha)$, arranged in increasing order of magnitude, then for any $\alpha > 0$ bounds for these zeros are given in* (33), *and also for any $\alpha > 0$*

$$(42)\qquad {}_Lx_{nn} = 4n + 2\alpha - 3.7115142\,(4n+2\alpha)^{1/3} + 2.7550676\,(4n+2\alpha)^{-1/3} + O(n^{-1}).$$

UNIVERSITY OF PENNSYLVANIA.

REMARKS ON THE PROBLEM OF PLATEAU

By E. F. Beckenbach

1. **Introduction.** We shall consider the problem of Plateau in the following form.

Problem of Plateau. *Given a Jordan curve* Γ *in xyz-space, determine functions* $x(u, v)$, $y(u, v)$, $z(u, v)$ *which are continuous for* $u^2 + v^2 \leqq 1$, *are harmonic and satisfy* $E = G$, $F = 0$, *where*

$$E = x_u^2 + y_u^2 + z_u^2, \quad F = x_u x_v + y_u y_v + z_u z_v, \quad G = x_v^2 + y_v^2 + z_v^2,$$

for $u^2 + v^2 < 1$, *and map* $u^2 + v^2 = 1$ *in a topological way on* Γ.

Any set of functions satisfying the above conditions are coördinate functions of a minimal surface bounded by Γ and given in isothermic representation.

The following theorems have been proved.

Theorem 1. *If* Γ *bounds some surface, of the type of the circular disc, with a finite area, then the problem of Plateau is solvable for* Γ.

Theorem 2. *The problem of Plateau is solvable for an arbitrary Jordan curve* Γ.

Theorem 1 has been proved separately and at about the same time by J. Douglas and T. Radó.[1] Subsequent proofs have been given by E. J. McShane[2] and R. Courant.[3] Theorem 2 has been proved by J. Douglas (loc. cit.), and later, by means of a different method but the same lemmas, by T. Radó.[4] In what follows, we consider alternative proofs of this latter theorem.

In proving Theorem 2, Douglas, assuming Theorem 1, first uses a limiting process to establish the existence of mapping functions. He then completes the proof by using the following two lemmas to show that the functions thus obtained map $u^2 + v^2 = 1$ topologically on Γ.

Lemma 1. *Let* $x(u, v)$, $y(u, v)$, $z(u, v)$ *be harmonic and satisfy* $E = G$, $F = 0$ *for* $u^2 + v^2 < 1$. *Suppose* $x(u, v)$, $y(u, v)$, $z(u, v)$ *remain continuous on an arc* σ *of* $u^2 + v^2 = 1$, *and* $x(u, v) = \text{const.} = x_0$, $y(u, v) = \text{const.} = y_0$, $z(u, v) = \text{const.} = z_0$ *on* σ. *Then* $x(u, v) \equiv x_0$, $y(u, v) \equiv y_0$, $z(u, v) \equiv z_0$.

Douglas' Lemma. *Let the integrable functions* $\xi(\varphi)$, $\eta(\varphi)$, $\zeta(\varphi)$, *substituted*

Received July 12, 1937.

[1] Their results are summed up in the following papers: J. Douglas, *Solution of the problem of Plateau,* Transactions of the American Mathematical Society, vol. 33 (1931), pp. 263–321; T. Radó, *The problem of the least area and the problem of Plateau,* Mathematische Zeitschrift, vol. 32 (1930), pp. 763–796.

[2] E. J. McShane, *Parametrization of saddle surfaces, with application to the problem of Plateau,* Transactions of the American Mathematical Society, vol. 35 (1933), pp. 716–733.

[3] R. Courant, *On the problem of Plateau,* Proceedings of the National Academy of Sciences, U. S. A., vol. 22 (1936), pp. 367–372.

[4] *An iterative process in the problem of Plateau,* Transactions of the American Mathematical Society, vol. 35 (1933), pp. 869–887.

in the Poisson integral formula, determine the (harmonic) coördinate functions of a minimal surface in isothermic representation. Let further $\xi(\varphi)$, $\eta(\varphi)$, $\zeta(\varphi)$ *approach definite limit values* $\xi_-(\pi)$, $\eta_-(\pi)$, $\zeta_-(\pi)$ *and* $\xi_+(\pi)$, $\eta_+(\pi)$, $\zeta_+(\pi)$ *according as* $\varphi \to \pi$ *in clockwise and counterclockwise senses, respectively. Then*

$$\xi_-(\pi) = \xi_+(\pi), \qquad \eta_-(\pi) = \eta_+(\pi), \qquad \zeta_-(\pi) = \zeta_+(\pi).$$

We have the following generalization[5] of a theorem of Lindelöf.[6]

LEMMA 2. *Let* $x(u, v)$, $y(u, v)$, $z(u, v)$ *be harmonic and bounded and satisfy* $E = G$, $F = 0$ *for* $0 < \text{arc tan}\,(v/u) < \alpha$, $0 < u^2 + v^2 < r_0^2$. *Suppose* $x(u, v)$, $y(u, v)$, $z(u, v)$ *remain continuous on the ray* $0 < u < r_0$, $v = 0$, *and* $x(u, 0) \to x_0$, $y(u, 0) \to y_0$, $z(u, 0) \to z_0$ *as* $u \to +0$. *Then in every sector*

$$0 < \text{arc tan}\,\frac{v}{u} < \alpha - \sigma, \qquad u^2 + v^2 < r_0^2, \qquad \textit{where } \sigma > 0,$$

we have $x(u, v) \to x_0$, $y(u, v) \to y_0$, $z(u, v) \to z_0$ *as* $(u, v) \to (0, 0)$ *in any manner.*

Since Douglas' lemma is a direct consequence[7] of Lemma 2, it follows, as Radó has remarked,[8] that Theorem 2 is a consequence of Theorem 1 and Lemmas 1 and 2.

We shall call attention to two pairs of proofs, most of which have previously been given, of Lemmas 1 and 2, and then, reviewing the limiting process, shall obtain a proof of Theorem 2 from Theorem 1 and Lemmas 1 and 2.

1. **Proof of Lemmas 1 and 2 by means of subharmonic functions.** A lemma which allows the immediate application of the Principle of the Maximum to minimal surfaces is the following.[9]

A necessary and sufficient condition that the continuous functions $x(u, v)$, $y(u, v)$, $z(u, v)$ *be harmonic functions satisfying* $E = G$, $F = 0$, *is that* $[(x - a)^2 + (y - b)^2 + (z - c)^2]^{\frac{1}{2}}$ *be of class*[10] *PL for arbitrary choice of the real constants* a, b, c.

By means of the above lemma, Beckenbach and Radó give brief proofs of Lemmas 1 and 2, strictly analogous to proofs by means of the Principle of the Maximum of corresponding theorems concerning analytic functions of a complex variable. Actually, they prove Lemma 1 under the additional restriction that $x(u, v)$, $y(u, v)$, $z(u, v)$ be bounded. But this restriction may be removed as follows. If $x(u, v)$, $y(u, v)$, $z(u, v)$ satisfy the assumptions of Lemma 1, then in an arbitrary Jordan region R bounded by $\sigma + B$, where $\sigma + B$ is a Jordan

[5] E. F. Beckenbach and T. Radó, *Subharmonic functions and minimal surfaces*, Transactions of the American Mathematical Society, vol. 35 (1933), pp. 648–661.

[6] See Pólya und Szegö, *Aufgaben und Lehrsätze aus der Analysis*, Berlin, 1925, vol. I, p. 138, problem 277.

[7] E. F. Beckenbach and T. Radó, loc. cit., p. 658.

[8] T. Radó, *On the problem of Plateau*, Berlin, 1933, p. 73.

[9] E. F. Beckenbach and T. Radó, loc. cit., p. 654.

[10] A function $p(u, v)$, defined in a domain D, is said to be of class PL in D provided $p(u, v)$ is continuous and $\geqq 0$ in D and $\log p(u, v)$ is subharmonic in the part of D where $p(u, v) > 0$.

curve and every point of B is in $u^2 + v^2 < 1$, the functions $x(u, v)$, $y(u, v)$, $z(u, v)$ are bounded. Map $\alpha^2 + \beta^2 < 1$ conformally on the interior of R by means of the analytic function $u + iv = f(\alpha + i\beta)$; an arc σ' of $\alpha^2 + \beta^2 = 1$ will correspond to σ. There are induced functions $x = x(u, v) = X(\alpha, \beta)$, etc., which satisfy the conditions, with α, β, σ' replacing u, v, σ, under which Beckenbach and Radó proved Lemma 1, so that $X(\alpha, \beta) \equiv x_0$, $Y(\alpha, \beta) \equiv y_0$, $Z(\alpha, \beta) \equiv z_0$, whence $x(u, v) \equiv x_0$, $y(u, v) \equiv y_0$, $z(u, v) \equiv z_0$.

2. **Alternative proofs of Lemmas 1 and 2.** Radó has proved[11] Lemma 1 about as follows. Since by assumption the harmonic functions $x(u, v)$, $y(u, v)$, $z(u, v)$ reduce to constants on σ, it follows by the Principle of Symmetry that these functions remain analytic on σ, and consequently the relations $E = G$, $F = 0$ hold on σ. Since, for an isothermic map, we have $dx^2 + dy^2 + dz^2 = \lambda(du^2 + dv^2)$, where $\lambda = E = G$, and since $dx^2 + dy^2 + dz^2 = 0$ for (u, v) on σ, it follows that $E = G = 0$ on σ, and consequently $x_u = x_v = y_u = y_v = z_u = z_v = 0$ on σ. That is, the functions $x_u - ix_v$, $y_u - iy_v$, $z_u - iz_v$, which are analytic functions of $w = u + iv$, vanish on an arc of their domain of regularity and therefore vanish identically. Hence, $x_u \equiv x_v \equiv y_u \equiv y_v \equiv z_u \equiv z_v \equiv 0$, so that $x(u, v)$, $y(u, v)$, $z(u, v)$ are identically constant.

We now offer a companion proof, based on the notion of normal families, of Lemma 2.

Three functions $x(u, v)$, $y(u, v)$, $z(u, v)$, harmonic in a domain D, are called a *triple of conjugate harmonic functions* provided they satisfy $E = G$, $F = 0$ in D. In conformity with analytic function theory, we shall say that a family of such triples constitutes a *normal family of triples of conjugate harmonic functions* in D provided that every infinite sequence of triples of the family contains a subsequence of triples which converges uniformly to a triple of conjugate harmonic functions, or for which $x^2 + y^2 + z^2$ converges uniformly to infinity, in every closed region in D. It is a well known fact[12] that a family of functions $\{h(u, v)\}$, harmonic and uniformly bounded in D, constitutes a normal family of harmonic functions in D, and that if

$$[h_n(u, v)], \qquad n = 0, 1, 2, \cdots,$$

is a convergent sequence of the family, then the sequence

$$\left[\frac{\partial^{j+k}}{\partial u^j \partial v^k} h_n(u, v)\right], \qquad n = 0, 1, 2, \cdots,$$

j and k being fixed, converges uniformly in every closed region in D to the corresponding derivative of the limit function. It follows immediately that a family of triples of conjugate harmonic functions, uniformly bounded in a domain D, constitutes a normal family of triples of conjugate harmonic functions.

[11] T. Radó, *Some remarks on the problem of Plateau*, Proceedings of the National Academy of Sciences, U. S. A., vol. 16 (1930), pp. 242–248.

[12] See O. D. Kellogg, *Foundations of Potential Theory*, Berlin, 1929, Chapter X.

Denote[13] $(u, v) = (0, 0)$ by O, $(r_0, 0)$ by A, and $(r_0 \cos \alpha, r_0 \sin \alpha)$ by B. Draw a line through O making an angle $\alpha - \sigma > 0$ with OA and cutting the arc AB at C. Let $0 < \delta < r_0$. Construct arcs with center O, radii $\delta/2^n$, cutting OA at A_n, and OC at C_n, $n = 0, 1, 2, \cdots$. Let D_n be the domain bounded by $\overline{A_n A_{n+1}} \overset{\frown}{C_{n+1}} \overline{C_{n+1} C_n} \overset{\frown}{A_n}$. Then, for (u, v) in D_0, $(u/2^n, v/2^n)$ is in D_n.

Define

$$x_n(u, v) = x\left(\frac{u}{2^n}, \frac{v}{2^n}\right), \qquad y_n(u, v) = y\left(\frac{u}{2^n}, \frac{v}{2^n}\right), \qquad z_n(u, v) = z\left(\frac{u}{2^n}, \frac{v}{2^n}\right).$$

In D_0, $x_n(u, v)$, $y_n(u, v)$, $z_n(u, v)$ take on the same values that $x(u, v)$, $y(u, v)$, $z(u, v)$ take on in D_n. Since $x(u, v)$, $y(u, v)$, $z(u, v)$ are bounded, the sequence

$$(1) \qquad [x_n(u, v), y_n(u, v), z_n(u, v)], \qquad n = 0, 1, 2, \cdots,$$

forms a normal family of triples of conjugate harmonic functions in D_0. Then there is a subsequence

$$[x_{n_k}(u, v), y_{n_k}(u, v), z_{n_k}(u, v)], \qquad k = 0, 1, 2, \cdots,$$

which converges uniformly in D_0 to a set of conjugate harmonic functions $\bar{x}(u, v)$, $\bar{y}(u, v)$, $\bar{z}(u, v)$.

Since $x(u, v)$ is continuous on OA, $x_{n_k}(u, v)$ is continuous there, and, for $(u, 0)$ on the boundary of D_0,

$$\bar{x}(u, 0) = \lim_{k \to \infty} x_{n_k}(u, 0) = \lim_{k \to \infty} x\left(\frac{u}{2^{n_k}}, 0\right) = x_0;$$

similarly, $\bar{y}(u, 0) = y_0$, $\bar{z}(u, 0) = z_0$. Therefore, by Lemma 1, $\bar{x}(u, v) \equiv x_0$, $\bar{y}(u, v) \equiv y_0$, $\bar{z}(u, v) \equiv z_0$.

Now the entire sequence (1) must converge to x_0, y_0, z_0, since otherwise there would be a subsequence which converges to a triple of conjugate harmonic functions other than x_0, y_0, z_0, and the above analysis shows that any convergent subsequence converges to x_0, y_0, z_0. Therefore, for (u, v) in D_0,

$$\lim_{n \to \infty} x_n(u, v) = x_0, \qquad \lim_{n \to \infty} y_n(u, v) = y_0, \qquad \lim_{n \to \infty} z_n(u, v) = z_0;$$

but the values of $x_n(u, v)$, $y_n(u, v)$, $z_n(u, v)$ in D_0 are the values of $x(u, v)$, $y(u, v)$, $z(u, v)$ in D_n, so that $x(u, v) \to x_0$, $y(u, v) \to y_0$, $z(u, v) \to z_0$ as $(u, v) \to (0, 0)$ in the sector $0 < \arctan (v/u) < \alpha - \sigma$.

3. **Proof of Theorem 2.** Approximate to Γ in the sense of Fréchet by a sequence Γ_n, $n = 0, 1, 2, \cdots$, of simple closed polygons. By Theorem 1, the problem of Plateau is solvable for Γ_n; further, by means of an adjoined linear fractional transformation, the solution can be so normalized that three distinct points A, B, C on $u^2 + v^2 = 1$ are carried into three arbitrary distinct points

[13] The following proof parallels a proof of Montel for analytic functions. See P. Montel, *Leçons sur les familles normales de fonctions analytiques*, Paris, 1927, pp. 188–192

A_n, B_n, C_n on Γ_n. Choose three distinct points A^*, B^*, C^* on Γ and let $A_n \to A^*$, $B_n \to B^*$, $C_n \to C^*$. Let now

$$x = x_n(u, v), \qquad y = y_n(u, v), \qquad z = z_n(u, v)$$

solve the normalized problem for Γ_n, and let the corresponding boundary functions be

$$x = \xi_n(\varphi), \qquad y = \eta_n(\varphi), \qquad z = \zeta_n(\varphi), \qquad \varphi = \text{arc tan}\,\frac{v}{u}.$$

An immediate generalization[14] of the fact that a uniformly bounded sequence of monotonic functions must contain a convergent subsequence assures us of the existence of a subsequence

$$(2) \qquad x = \xi_{n_k}(\varphi), \qquad y = \eta_{n_k}(\varphi), \qquad z = \zeta_{n_k}(\varphi), \qquad k = 0, 1, 2, \cdots,$$

converging everywhere on $u^2 + v^2 = 1$ to limit functions

$$(3) \qquad x = \xi(\varphi), \qquad y = \eta(\varphi), \qquad z = \zeta(\varphi),$$

which map $u^2 + v^2 = 1$ monotonically on Γ, carrying A, B, C to A^*, B^*, C^* respectively.

Consider the harmonic functions

$$(4) \quad x = x(u, v), \qquad y = y(u, v), \qquad z = z(u, v), \qquad u^2 + v^2 < 1,$$

obtained by substituting the functions (3) in the Poisson integral formula. Since the functions (2) are uniformly bounded, we may pass to the limit in the corresponding sequences of Poisson integrals, so that the functions

$$x_{n_k}(u, v), \qquad y_{n_k}(u, v), \qquad z_{n_k}(u, v), \qquad k = 0, 1, 2, \cdots,$$

and their partial derivatives converge in $u^2 + v^2 < 1$ to the functions (4) and their corresponding partial derivatives. Since $E_{n_k} = G_{n_k}$, $F_{n_k} = 0$, it follows therefore that $E = G$, $F = 0$.

That the conjugate harmonic functions (4) give a solution of the problem of Plateau for the curve Γ will be established when we show that the boundary functions (3), which map $u^2 + v^2 = 1$ monotonically on Γ, actually map $u^2 + v^2 = 1$ topologically on Γ. Now the functions (3) cannot remain constant on an arc of $u^2 + v^2 = 1$, by Lemma 1 and the three-point condition. And, by the monotonic character of the map, the functions (3) have definite one-sided limits for each value of φ; we shall show that these limits are the same from both sides at an arbitrary φ_0. Because of the nature of the possible discontinuity of $\xi(\varphi)$ at $\varphi = \varphi_0$, it follows from a well-known property of the Poisson integral that $x(u, v)$ approaches a definite limit if $(u, v) \to (\cos \varphi_0, \sin \varphi_0)$ along any straight line in $u^2 + v^2 < 1$, this limit being a linear function of the angle which the straight line makes with a fixed direction and varying from $\xi_-(\varphi_0)$ to $\xi_+(\varphi_0)$. Similar statements hold for $y(u, v)$ and $z(u, v)$. But if

[14] T. Radó, first footnote, p. 771.

we join two such straight lines by a circular arc lying in $u^2 + v^2 < 1$, we obtain a sector for which Lemma 2 applies; consequently, $(x, y, z) \to$ a definite (x_0, y_0, z_0) which does not vary with the angle. That is, the linear functions mentioned above are constants, whence

$$\xi_-(\varphi_0) = \xi_+(\varphi_0), \qquad \eta_-(\varphi_0) = \eta_+(\varphi_0), \qquad \zeta_-(\varphi_0) = \zeta_+(\varphi_0).$$

Therefore, the functions (3) map $u^2 + v^2 = 1$ in a one-to-one way on Γ.

4. Lemmas 1 and 2 are essentially theorems *im kleinen*, and their proofs are independent of the dimensionality of the containing Euclidean space. Therefore, these lemmas may be used to discuss the behavior on the boundary of functions giving isothermic representations of minimal surfaces bounded by several Jordan curves in Euclidean n-space.

RICE INSTITUTE.

ANALYTIC FUNCTIONS OF ABSOLUTELY CONVERGENT GENERALIZED TRIGONOMETRIC SUMS

By R. H. Cameron

1. **Introduction.** It has been shown by Wiener[1] that a nowhere vanishing periodic function with an absolutely convergent Fourier series has a reciprocal whose Fourier series also converges absolutely. Lévy[2] has pointed out that this result can be extended from reciprocals to general analytic functions. Thus if $f(x)$ is periodic and never zero and has an absolutely convergent Fourier series, it follows that $F[f(x)]$ also has an absolutely convergent Fourier series provided that $F(z)$ is analytic and single valued whenever $z = f(x)$. One of the results of this paper (Theorem I) shows that these results are true in n or even $\aleph_0$ dimensions. This is accomplished by carrying through Wiener's proof with the necessary modifications to take care of dimensionality.

One might reasonably ask whether this result can be extended from periodic to almost periodic functions. A partial answer to this question has been given by Bochner,[3] who has shown that reciprocals of trigonometric polynomials which are bounded away from zero on the real axis have absolutely convergent Fourier series. It is shown in the present paper[4] that the theorem is true not only for trigonometric polynomials, but also for absolutely convergent infinite trigonometric sums. No further hypothesis is required; so the exponents are altogether unrestricted and may be any countable set of real numbers. Moreover this result is true not only for reciprocals, but for all analytic functions; and it holds in n or even $\aleph_0$ dimensions. Thus the final result of the paper is

Theorem II. *Let $f(x_1, x_2, \cdots)$ be an almost periodic function with an absolutely convergent Fourier series, and let R be the closure of its set of values. Then if $F(z)$ is a function analytic over an open set S containing R, it follows that $F[f(x_1, x_2, \cdots)]$ is an almost periodic function with an absolutely convergent Fourier series.*

Received July 12, 1937.

[1] N. Wiener, *Tauberian theorems*, Ann. of Math., (2), vol. 33 (1932), pp. 1–100; p. 14.

[2] P. Lévy, *Sur la convergence absolue des séries de Fourier*, C. R. Acad. Sci., Paris, vol. 196 (1933), pp. 463–464.

[3] S. Bochner, *Beitrag zur absoluten Konvergenz fastperiodischer Fourierreihen*, Jahresbericht der Deutschen Math. Ver., vol. 39 (1930), pp. 52–54.

[4] After this paper had been submitted for publication, the author learned that his main theorem (without the extension to analytic functions or to more than one dimension) has been proved independently by H. R. Pitt. Apparently Pitt's work was done somewhat earlier than the author's, though it was not submitted for publication until about the time the present paper was accepted for publication. It will appear in an early issue of the Journal of Mathematics and Physics, Massachusetts Institute of Technology.

2. **Absolute convergence a local property for periodic functions.** Before proving Theorem I we shall need to extend to infinitely many variables Wiener's lemma[5] that a periodic function has an absolutely convergent Fourier series if in the neighborhood of every point it is equal to a function having an absolutely convergent Fourier series. Such a generalization naturally depends on the type of neighborhoods we use, and the appropriate neighborhoods in this case are defined as follows. We consider the space whose points $P(x_1, x_2, \cdots)$ are unrestricted sequences $x_1, x_2, \cdots$ of real numbers. Then corresponding to each point $P(x_1, x_2, \cdots)$, each $\epsilon > 0$ and each positive integer n we define the (ϵ, n)-neighborhood of P to be the set of all points $x_1', x_2', \cdots$ satisfying

$$|x_j' - x_j| < \epsilon \pmod{2\pi} \qquad (j = 1, \cdots, n).$$

Jessen[6] has shown that for such neighborhoods in which all but the first n variables are unrestricted the Heine-Borel theorem holds for the whole space. This fact enables us to extend Wiener's proof to infinitely many variables and obtain

LEMMA 1. *Let $f(P) = f(x_1, x_2, \cdots)$ be periodic of period 2π in each variable. Suppose further it is known that corresponding to each point $P'(x_1', x_2', \cdots)$ there exist $\epsilon_{P'} > 0$ and $n_{P'}$ and a function $f_{P'}(P) = f_{P'}(x_1, x_2, \cdots)$ which equals $f(P)$ throughout the $(\epsilon_{P'}, n_{P'})$-neighborhood of P' and has an absolutely convergent Fourier series*

$$f_{P'}(P) = A_0^{(P')} + \sum_{n=1}^{\infty} A_n^{(P')} \exp\left\{i \sum_{j=1}^{k_n^{(P')}} p_{n,j}^{(P')} x_j\right\}.$$

Then it follows that $f(P)$ has an absolutely convergent Fourier series

$$f(P) = A_0 + \sum_{n=1}^{\infty} A_n \exp\left\{i \sum_{j=1}^{k_n} p_{n,j} x_j\right\}.$$

For by the Heine-Borel theorem there are a finite number of points $P_1, P_2, \cdots, P_\nu$ such that every point P is contained in $N_{P_1} + N_{P_2} + \cdots + N_{P_\nu}$, where N_{P_j} is the $(\frac{1}{2}\epsilon_{P_j}, n_{P_j})$-neighborhood of P_j; and we shall show how to fit together absolutely convergent Fourier series in these neighborhoods to make $f(P)$. For positive values of $\xi < \pi$, let $T_\xi(x)$ be periodic of period 2π in x, and let it be defined by the equation $T_\xi(x) = \max\,[1 - |x|/\xi, 0]$ in one period $-\pi \leqq x \leqq \pi$.

Obviously this function consists of equally spaced isosceles peaks of height 1 with horizontal lines of height zero in between. Again, let

$$T_{\xi,m}(x_1, x_2, \cdots) = T_\xi(x_1)T_\xi(x_2) \cdots T_\xi(x_m),$$

and note that $T_{\xi,m}(x_1, x_2, \cdots)$ vanishes outside of the (ξ, m)-neighborhood of the origin and has a peak (or infinite dimensional edge) of unit height a·

[5] N. Wiener, loc. cit., p. 10.

[6] Jessen, *The theory of integration in a space of an infinite number of dimensions*, Acta Math., vol. 63 (1934), pp. 249–323; p. 256.

$x_1 = 0,\ x_2 = 0, \cdots, x_m = 0$. Finally, if $\epsilon = \min(\epsilon_{P_1}, \cdots, \epsilon_{P_\nu})$ and $n = \max(n_{P_1}, \cdots, n_{P_\nu})$, and if λ is an integer so great that $2^{-\lambda} < \epsilon/(2\pi)$, let

$$U_{\mu_1,\cdots,\mu_n}(P) = U_{\mu_1,\cdots,\mu_n}(x_1, x_2, \cdots) = T_{2^{-\lambda}\pi, n}(x_1 - 2^{-\lambda}\pi\mu_1, \cdots, x_n - 2^{-\lambda}\pi\mu_n)$$

and note that

$$\sum_{\mu_1=\cdots=\mu_n=0}^{2^{\lambda+1}-1} U_{\mu_1,\cdots,\mu_n}(P) \equiv 1,$$

and hence that for all P,

$$f(P) = \sum_{\mu_1,\cdots,\mu_n=0}^{2^{\lambda+1}-1} U_{\mu_1,\cdots,\mu_n}(P)\, f(P). \tag{1}$$

Now if in each term of this sum we replace $f(P)$ by a function which equals $f(P)$ except when the coefficient of $f(P)$ is zero, the equation will still be true. But such functions can be found with absolutely convergent Fourier series. For

$$Q:\ 2^{-\lambda}\pi\mu_1,\ 2^{-\lambda}\pi\mu_2, \cdots, 2^{-\lambda}\pi\mu_n, 0, 0, \cdots$$

is contained in one of the neighborhoods $N_{P_1}, \cdots, N_{P_\nu}$, say N_{P_σ}; and it follows that the $(2^{-\lambda}\pi, n)$-neighborhood of Q is contained in the $(\epsilon_{P_\sigma}, n_{P_\sigma})$-neighborhood of P_σ. Thus for all P

$$U_{\mu_1,\cdots,\mu_n}(P) f(P) \equiv U_{\mu_1,\cdots,\mu_n}(P) f_{P_\sigma}(P);$$

and since $T_\xi(x)$ has an absolutely convergent Fourier series, so do $T_\xi(x_1, x_2, \cdots)$ and $U_{\mu_1,\cdots,\mu_n}(P)$ and $U_{\mu_1,\cdots,\mu_n}(P)f(P)$. Consequently, it follows from (1) that $f(P)$ has an absolutely convergent Fourier series and the lemma holds.

3. Fourier series of small absolute value sum. Again following the course of the Wiener argument, we prove the

LEMMA 2. *Let $f(x_1, x_2, \cdots)$ have period 2π in each variable, and let it have an absolutely convergent Fourier series, the sum of the absolute values of the coefficients other than the constant term being K. Then if $F(z)$ is a function analytic inside and on the boundary of the circle $|z - f(0, 0, \cdots)| \leqq 2K$, it follows that $F[f(x_1, x_2, \cdots)]$ has an absolutely convergent Fourier series.*

For $F(z)$ can be expanded in a power series about $z_0 = f(0, 0, \cdots)$, and the Fourier series of $f(x_1, x_2, \cdots)$ can be formally substituted for z in this power series. The sum of the absolute values of the terms arising from $(z - z_0)^n$ will be less than or equal to $(2K)^n$, and hence the whole series with all parentheses removed will be absolutely convergent.

4. General periodic functions with absolutely convergent Fourier series. We are now in a position to prove

THEOREM I. *Let $f(x_1, x_2, \cdots)$ be a function of period 2π in each variable, and let the sum $\sum_{n=0}^{\infty} A_n$ of the coefficients of its Fourier series*

$$f(x_1, x_2, \cdots) = A_0 + \sum_{n=1}^{\infty} A_n \exp\left[i \sum_{j=1}^{k_n} p_{n,j} x_j\right] \tag{2}$$

be absolutely convergent. Then if R is the closure of the range of $f(x_1, x_2, \cdots)$ and the function $F(z)$ is analytic in an open set S containing R, it follows that $F[f(x_1, x_2, \cdots)]$ also has an absolutely convergent series:

$$F[f(x_1, x_2, \cdots)] = B_0 + \sum_{n=1}^{\infty} B_n \exp\left[i \sum_{j=1}^{l_n} q_{n,j} x_j\right]. \tag{3}$$

On account of Lemma 1, we need only show that corresponding to each point $P'(x_1', x_2', \cdots)$ there exists an $\epsilon_{P'} > 0$ and a positive integer $n_{P'}$ and a function $g_{P'}(x_1, x_2, \cdots)$ which has an absolutely convergent Fourier series and which equals $F[f(x_1, x_2, \cdots)]$ throughout the $(\epsilon_{P'}, n_{P'})$-neighborhood of P'. For under these circumstances Lemma 1 establishes the existence and absolute convergence of the series (3). And since the function $f^*_{P'}(x_1, x_2, \cdots) = f(x_1 - x_1', x_2 - x_2', \cdots)$ satisfies the same hypothesis as $f(x_1, x_2, \cdots)$, it follows that we need only consider the origin and show that there is a function $g_0(x_1, x_2, \cdots)$ which equals $F[f(x_1, x_2, \cdots)]$ throughout some neighborhood of the origin and has an absolutely convergent Fourier series.

Now for any function g let $\sigma(g)$ denote the sum of the absolute values of the Fourier series of the function g, and let $f(0, 0, \cdots) = z_0$, so that

$$f(x_1, x_2, \cdots) - z_0 = \sum_{n=1}^{\infty} A_n \left\{\exp\left[i \sum_{j=1}^{k_n} p_{n,j} x_j\right] - 1\right\}.$$

Let $\delta > 0$ be so small that z is in S if $|z - z_0| \leqq \delta$, and let N be so great that

$$\sum_{n=N+1}^{\infty} |A_n| < \frac{\delta}{8}.$$

Thus if

$$g(P) = g(x_1, x_2, \cdots) = \sum_{n=1}^{N} A_n \left\{\exp\left[i \sum_{j=1}^{k_n} p_{n,j} x_j\right] - 1\right\}$$

and

$$h(P) = \sum_{n=N+1}^{\infty} A_n \left\{\exp\left[i \sum_{j=1}^{k_n} p_{n,j} x_j\right] - 1\right\},$$

it follows that $f(P) = z_0 + g(P) + h(P)$ so that $\sigma(h) < \frac{1}{4}\delta$. Now for $0 < \xi < \frac{1}{2}\pi$ define $W_\xi(P) = \prod_{n=1}^{N} V_\xi(x_n)$, where

$$V_\xi(x) = \begin{cases} 0 & \text{if} \quad |x| \geqq 2\xi \pmod{2\pi}, \\ 2 - \dfrac{|x|}{\xi} & \text{if} \quad \xi \leqq |x| \leqq 2\xi \pmod{2\pi}, \\ 1 & \text{if} \quad |x| \leqq \xi \pmod{2\pi}; \end{cases}$$

and note by actual computation that $V_\xi(x)$ (and hence also $W_\xi(P)$) has an absolutely convergent Fourier series whose absolute value sum is a bounded function of ξ. Moreover, $\lim_{\xi \to 0^+} \sigma[V_\xi(x)(e^{ipx} - 1)] = 0$; for if Σ' denotes the

sum from $n = -\infty$ to $n = +\infty$, omitting 0 and p, we have by actual computation for all integers p,

$$
\begin{aligned}
&\sigma[V_\xi(x)(e^{ipx} - 1)] \\
&\quad = 2\left|\frac{3\xi}{2\pi} - \frac{\cos p\xi - \cos 2p\xi}{p^2\pi\xi}\right| \\
&\qquad + {\sum}'\left|\frac{2[\sin\frac{1}{2}(2n-p)\xi \sin\frac{1}{2}p\xi - \sin(2n-p)\xi\sin p\xi]}{(n-p)^2\pi\xi}\right. \\
&\qquad\qquad \left. + \frac{2(2np - p^2)\sin\frac{3}{2}n\xi \sin\frac{1}{2}n\xi}{(n-p)^2 n^2\pi\xi}\right| \\
&\quad \leqq 2\left|\frac{3\xi}{2\pi} - \frac{2\sin\frac{3}{2}p\xi\sin\frac{1}{2}p\xi}{p^2\pi\xi}\right| \\
&\qquad + {\sum}'\left\{\frac{2\,|\sin\frac{1}{2}(2n-p)\xi|^{\frac{1}{2}}\cdot|\sin\frac{1}{2}p\xi| + 2\,|\sin(2n-p)\xi|^{\frac{1}{2}}\cdot|\sin p\xi|}{(n-p)^2\pi\xi}\right. \\
&\qquad\qquad \left. + \frac{2\,|2np - p^2|\cdot|\sin\frac{3}{2}n\xi|\cdot|\sin\frac{1}{2}n\xi|^{\frac{1}{2}}}{(n-p)^2 n^2\pi\xi}\right\} \\
&\quad \leqq 2\left|\frac{3\xi}{2\pi}\right| + 2\left|\frac{2\,\frac{3}{2}p\xi\cdot\frac{1}{2}p\xi}{p^2\pi\xi}\right| + {\sum}'\left\{\frac{2\,|\frac{1}{2}(2n-p)\xi|^{\frac{1}{2}}\cdot|\frac{1}{2}p\xi| + 2\,|(2n-p)\xi|^{\frac{1}{2}}\cdot|p\xi|}{(n-p)^2\pi\xi}\right. \\
&\qquad\qquad \left. + \frac{2\,|2np - p^2|\cdot|\frac{3}{2}n\xi|\cdot|\frac{1}{2}n\xi|^{\frac{1}{2}}}{(n-p)^2 n^2\pi\xi}\right\} \\
&\quad \leqq \frac{6\xi}{\pi} + \xi^{\frac{1}{2}}{\sum}'\left\{\frac{p\,|\frac{1}{2}(2n-p)|^{\frac{1}{2}} + 2p\,|2n-p|^{\frac{1}{2}}}{(n-p)^2\pi} + \frac{2^{\frac{1}{2}}\,|2np - p^2|\cdot\frac{3}{2}}{(n-p)^2|n|^{\frac{1}{2}}\pi}\right\}.
\end{aligned}
$$

Now consider $W_\xi(P)g(P)$, and note that

$$\lim_{\xi\to 0^+}\sigma[W_\xi(P)g(P)] = 0,$$

since

$$\lim_{\xi\to 0^+}\sigma[W_\xi(P)e^{i(p_{m+1}x_{m+1}+\cdots+p_Nx_N)}(e^{ip_mx_m} - 1)] = 0$$

holds for any integers p_m and implies

$$\lim_{\xi\to 0^+}\sigma[W_\xi(P)(e^{i(p_1x_1+p_2x_2+\cdots+p_Nx_N)} - 1)] = 0.$$

Finally, choose $\epsilon > 0$ so small that

$$\sigma[W_\epsilon(P)g(P)] < \tfrac{1}{4}\delta,$$

and consider the functions

$$\bar{f}(P) = z_0 + W_\epsilon(P)g(P) + h(P).$$

Since

$$g(0, 0, \cdots) = h(0, 0, \cdots) = 0,$$

the constant term of $\bar{f}(P)$ is $\bar{f}(0, 0, \cdots) = z_0$; and since

$$\sigma[\bar{f}(P) - z_0] = \sigma[W_\epsilon(P)g(P) + h(P)] < \tfrac{1}{2}\delta,$$

Lemma 2 applies to $\bar{f}(P)$ and shows that $g_0(P) = F[\bar{f}(P)]$ has an absolutely convergent Fourier series. But by definition $W_\epsilon(P) \equiv 1$ throughout the (ϵ, N)-neighborhood of the origin, and hence $g_0(P) = F[f(P)]$ in the neighborhood, and our proof is complete.

5. **Almost periodic functions.** We can now pass to the case of generalized Fourier series and prove Theorem II, which has been stated in the Introduction. Let

$$f(x_1, x_2, \cdots) = A_0 + \sum_{n=1}^{\infty} A_n \exp\left[i \sum_{j=1}^{p_n} \lambda_{n,j} x_j\right].$$

The proof can be based on Theorem I in the following way. Let $\delta > 0$ be so small that if $|z_1 - z_2| \leqq \delta$ and $z_1 \epsilon R_1$, then $z_2 \epsilon S$; and let R^* be the closed set consisting of all points whose distance from R is not greater than δ. Let N be so great that $\sum_{n=N+1}^{\infty} |A_n| < \tfrac{1}{2}\delta$, and let $\mu_1, \mu_2, \cdots, \mu_s$ be an integral basis for all $\lambda_{n,j}$ for which $n \leqq N$ and $j \leqq p_n$; so that

$$\lambda_{n,j} = \sum_{\nu=1}^{s} k_{n,j,\nu}\mu_\nu \qquad (n = 1, \cdots, N; j = 1, \cdots, p_n).$$

Here the $k_{n,j,\nu}$ are integers, and no integers $k_1, \cdots, k_s$ except $0, \cdots, 0$ make $\sum_{\nu=1}^{s} k_\nu \mu_\nu = 0$.

If p is the greatest of $p_1, \cdots, p_N$, consider the functions[7]

$$h(Y_{1,1}, \cdots, Y_{1,s}; \cdots; Y_{p,1}, \cdots, Y_{p,s}; \xi_{N+1}, \xi_{N+2}, \cdots) = A_0 + \sum_{n=1}^{N} A_n \exp\left\{i \sum_{j=1}^{p_n} \sum_{\nu=1}^{s} k_{n,j,\nu} Y_{j,\nu}\right\}$$

and

$$g(Y_{1,1}, \cdots, Y_{1,s}; \cdots; Y_{p,1}, \cdots, Y_{p,s}; \xi_{N+1}, \xi_{N+2}, \cdots) = h(Y_{1,1}, \cdots, Y_{1,s}; \cdots; Y_{p,1}, \cdots, Y_{p,s}; \xi_{N+1}, \xi_{N+2}, \cdots) + \sum_{n=N+1}^{\infty} A_n e^{i\xi_n}.$$

It is clear that g has period 2π in all its arguments, and that

$$g\left(\mu_1 x_1, \cdots, \mu_s x_1; \cdots; \mu_1 x_p, \cdots, \mu_s x_p; \sum_{j=1}^{p_{N+1}} \lambda_{N+1,j} x_j, \sum_{j=1}^{p_{N+2}} \lambda_{N+2,j} x_j, \cdots\right) = f(x_1, x_2, \cdots).$$

[7] Putting in exponents arbitrarily as we have done after the N-th term changes the range of the function; while putting them in according to the basis as Bochner did in his paper (loc. cit.) and as we have done in the earlier terms would require the use of limit periodic functions if it were carried out for all the terms. The author wishes to thank Professor Norbert Wiener for suggesting this combination of the two methods.

Moreover, the closure of the range of h is the same as the closure of the range of

$$A_0 + \sum_{n=1}^{N} A_n \exp i \sum_{j=1}^{p_n} \lambda_{n,j} x_j .$$

Thus by Theorem I it follows that $F(g)$ has an absolutely convergent Fourier series

$$B_0 + \sum_{n=1}^{\infty} B_n \exp \left[i \sum_{j=1}^{p} \sum_{\nu=1}^{s} q_{n,j,\nu} Y_{j,\nu} + i \sum_{j=N+1}^{l_n} r_{n,j} \xi_j \right].$$

Hence

$$F[f(x_1, x_2, \cdots)] = B_0 + \sum_{n=1}^{\infty} B_n \exp \left[i \sum_{j=1}^{p} \sum_{\nu=1}^{s} q_{n,j,\nu} \mu_\nu x_j + i \sum_{j=N+1}^{l_n} \sum_{\nu=1}^{p_j} r_{n,j} \lambda_{j,\nu} x_\nu \right],$$

and Theorem II is proved.

MASSACHUSETTS INSTITUTE OF TECHNOLOGY.

TRANSFORMATIONS ON SEQUENCE SPACES

By L. W. Cohen and Nelson Dunford

Hardy and Littlewood,[1] Littlewood[2] and others have given certain necessary and other sufficient conditions on the matrix a_{ij} in order that the bilinear form $\Sigma\Sigma a_{ij}x_ix_j$ be bounded for $\Sigma \mid x_i \mid^p \leqq 1$, $\Sigma \mid y_i \mid^q \leqq 1$. So far as we know no conditions on the matrix a_{ij} alone have been given which are necessary as well as sufficient for the boundedness of the corresponding bilinear form. In this paper we consider among other questions the more precise problem of determining the norm of the linear transformation $y = Tx$ on l_p to l_q in terms of the elements a_{ij} of the matrix representing this transformation. We have been successful in the special cases where T is on l_1 to l_p or c_0 and, less trivially, on l_p or c_0 to l_1, if $a_{ij} \geqq 0$. Conditions for the absolute convergence of the determinant of $(\delta_{ij} + a_{ij})$ representing $I + T$ as well as properties of the matrix of minors are also obtained. These last conditions together with necessary and sufficient conditions for compactness have been given for a Banach space with a denumerable basis φ_n. In such a space each element x is uniquely representable as

$$x = \sum_{n=1}^{\infty} x_n \varphi_n, \qquad x_n = T_n^{\Phi} x,$$

where T_n^{Φ} is a linear functional on the space Φ with $\mid T_n^{\Phi} \mid \leqq M_{\Phi}$. In such a space the convergence of x^m to x implies the uniform convergence of x_n^m to x_n.

In view of the well-known theorems on uniform boundedness of sequences of linear operations and the known conditions for weak convergence in many Banach spaces, it is comparatively trivial to give the form and norm of the general linear operation with the range in c, $m = l_\infty$, C, M (bounded functions), etc. Consequently these cases have been omitted from the discussion of such questions.

Theorem 1. *If Φ and Ψ are Banach spaces with denumerable bases and $Tx = y$ is a linear transformation of Φ into Ψ, the transformation is represented by*

$$y_i = \sum_{j=1}^{\infty} a_{ij} x_j, \tag{1}$$

Received June 2, 1937; presented to the American Mathematical Society, December, 1936, under the titles: L. W. Cohen, *Transformations on spaces with denumerable basis*; Nelson Dunford, *Linear transformations of sequences*.

[1] G. H. Hardy and J. E. Littlewood, *Bilinear forms bounded in space* [p, q], Quarterly Journal of Math., (Oxford), vol. 5 (1934), pp. 241–254.

[2] J. E. Littlewood, *On bounded bilinear forms in an infinite number of variables*, Quarterly Journal of Math., (Oxford), vol. 1 (1930), pp. 164–174.

where

$$a_{ij} = T_i^{\Psi} T \varphi_j .$$

The i-th equation is a linear functional $L_i x$ on Φ such that

$$| L_i | \leqq | T | M_{\Psi} ,$$

and the series (1) converge uniformly with respect to i. Conversely, if a_{ij} is a matrix such that for every $x \in \Phi$ there is a $y \in \Psi$ satisfying

$$(2) \qquad T_i^{\Psi} y = \sum_{j=1}^{\infty} a_{ij} T_j^{\Phi} x,$$

the transformation is continuous.

Proof. For each j

$$T\varphi_j = \psi_j^* = \sum_{i=1}^{\infty} a_{ij} \psi_i ,$$

where ψ_i is the basis of Ψ. For $x \in \Phi$, let $x^{(n)} = \sum_{j=1}^{n} x_j \varphi_j$. Then

$$y^n = Tx^{(n)} = \sum_{j=1}^{n} x_j \psi_j^* = \sum_{i=1}^{\infty} \psi_i \sum_{j=1}^{n} a_{ij} x_j ,$$

$$y = Tx = \sum_{i=1}^{\infty} y_i \psi_i ,$$

and

$$\left| y_i - \sum_{j=1}^{n} a_{ij} x_j \right| \leqq M_{\Psi} \| y - y^n \| \leqq M_{\Psi} | T | \cdot \| x - x^{(n)} \|$$

imply the uniform convergence of the series (1) with respect to i. Since $| y_i | \leqq M_{\Psi} | T | \cdot \| x \|$, $| L_i | \leqq M_{\Psi} | T |$.

Now, conversely, for fixed i and n the function $\sum_{j=1}^{n} a_{ij} T_j^{\Phi} x$ is continuous in x and hence for fixed i its limit, i.e., the function

$$y_i = \sum_{j=1}^{\infty} a_{ij} T_j^{\Phi} x,$$

is also continuous in x. Thus

$$y^n = \sum_{i=1}^{n} \left(\sum_{j=1}^{\infty} a_{ij} T_j^{\Phi} x \right) \psi_i$$

as well as its limit y is a continuous function of x.

In specializing the general linear transformation $Tx = y$, one may require that it be completely continuous, i.e., that it carry a bounded set into a compact set. We state the following condition for compactness.

THEOREM 2. *A set $X \subset \Phi$ is compact in Φ if and only if X is bounded and*

$$\lim_n \sum_{j=1}^{n} x_j \varphi_j = x$$

uniformly for $x \in X$.

Proof. From a sequence (y^m) of points in X a subsequence (x^m) may be chosen such that, for all i,

$$\lim_m T_i^{\Phi} x^m = a_i .$$

Now

$$\lim_n \sum_{i=1}^{n} T_i^{\Phi} x^m \varphi_i = x^m$$

uniformly with respect to m and

$$\lim_m \sum_{i=1}^{n} T_i^{\Phi} x^m \varphi_i$$

exists for all n. Thus $\lim_m x^m$ exists and X is compact.

Conversely, suppose that X is compact in Φ. The sequence

$$f_n(x) = \sum_{i=1}^{n} x_i \varphi_i$$

of linear transformations is convergent, hence equi-uniformly continuous. For any $\epsilon > 0$ there is a $\delta_\epsilon > 0$ such that, for all n,

$$\| f_n(x) \| < \frac{\epsilon}{3}, \qquad \| x \| < \delta_\epsilon .$$

There are $x^1, \cdots, x^{n(\epsilon)}$ in X such that for $x \in X$ and some $i = 1, \cdots, n(\epsilon)$

$$\| x - x^i \| < \min \left(\frac{\epsilon}{3}, \delta_\epsilon \right).$$

There is an N_ϵ such that $n > N_\epsilon$ implies

$$\| f_n(x^i) - x^i \| < \frac{\epsilon}{3} \qquad (i = 1, \cdots, n(\epsilon)).$$

Now for $n > N_\epsilon$, $x \in X$ and some i, we have

$$\| f_n(x) - x \| \leqq \| f_n(x) - f_n(x^i) \| + \| f_n(x^i) - x^i \| + \| x^i - x \| < \epsilon.$$

COROLLARY 1. *In l_p $(1 \leqq p < \infty)$ a set X is compact if and only if X is bounded and*

$$\lim_n \sum_{j=n}^{\infty} | x_j |^p = 0$$

uniformly with respect to (x_j) in X.

COROLLARY 2. *In c (the space of convergent sequences (x_j)) a set X is compact if and only if it is bounded and $\lim_j x_j$ exists uniformly with respect to (x_j) in X.*

COROLLARY 3. *The general completely continuous linear regular transformation of c into c is*

$$y_i = a_i \lim_j x_j + \sum_{j=1}^{\infty} a_{ij} x_j,$$

where

$$\lim_i a_i = 1, \qquad \lim_i \sum_{j=1}^{\infty} |a_{ij}| = 0$$

and the norm of the transformation is

$$\sup_i \left\{ |a_i| + \sum_{j=1}^{\infty} |a_{ij}| \right\}.$$

Thus, while every Toeplitz matrix is continuous, none is completely continuous on c to c. We have the formulas

$$y_i = a_i \lim_n \sum_{j=1}^{\infty} b_{nj} x_j + \sum_{j=1}^{\infty} a_{ij} x_j,$$

$$y_i = a_i \operatorname{Lim} x_i + \sum_{j=1}^{\infty} a_{ij} x_j,$$

where (b_{nj}) is a Toeplitz matrix and Lim is a generalized limit defined for all bounded sequences. These formulas yield completely continuous regular summation methods whose domain of definition is the set of all bounded sequences.

From Theorem 2 we also have

THEOREM 3. *If T is linear and completely continuous on Φ to Ψ and Ψ has a bounded basis, then*

$$\lim_n \psi_j^{*(n)} = \psi_j^* = T\varphi_j$$

uniformly in j.

It will appear below (Theorems 12, 13) that this condition is sufficient for the complete continuity of T in the cases where $\Phi = l_1$, $\Psi = l_p$ $(1 \leqq p < \infty)$, $\Psi = c_0$.

To formulate sufficient conditions for complete continuity we may use the following theorem which is well known in case the transformations involved are linear.

THEOREM 4. *If $T_n x = y$ is completely continuous on a Banach space S to a Banach space S' for each n, and if for any $\epsilon > 0$ there is an n_ϵ such that*

$$\| Tx - T_n x \| \leqq \epsilon \| x \|, \qquad (n > n_\epsilon),$$

then Tx is completely continuous.

Proof. Let X be a bounded subset of S and y^m a sequence in TX. There

is a sequence of sequences x_q^n in X such that each Tx_q^n is a y^m, $(x_q^n) \supset (x_q^{n+1})$ and $\lim_q T_n x_q^n$ exists. Writing $x^n = x_n^n$ we have for $\epsilon > 0$

$$\| Tx^n - Tx^{n+k} \| \leqq \| Tx^n - T_m x^n \| + \| T_m x^n - T_m x^{n+k} \|$$
$$+ \| T_m x^{n+k} - Tx^{n+k} \| < 2\epsilon M + \| T_m x^n - T_m x^{n+k} \|,$$

if $\| x \| \leqq M$ when $x \epsilon X$ and $m > m_\epsilon$. Then for $n > n_{m,\epsilon}$, $k > 0$ we have

$$\| Tx^n - Tx^{n+k} \| < (2M + 1)\epsilon.$$

Since S' is complete, TX is compact.

In case T is a linear transformation on Φ to Ψ, we have

THEOREM 5. *A necessary and sufficient condition for the complete continuity of T is that $\lim_n | T_n | = 0$, where $T_n x = y$ is defined by the matrix*

$$a_{ij}^{(n)} = 0, \qquad i < n,$$
$$= a_{ij}, \qquad i \geqq n,$$

associated with the matrix (a_{ij}) of T.

Proof. The sufficiency follows from the complete continuity of $T - T_n$, $\lim_n | T - (T - T_n) | = 0$ and Theorem 4. Conversely, let $y = Tx$ be completely continuous and

$$y = \sum_{i=1}^{\infty} y_i \psi_i, \qquad y^{(n)} = \sum_{i=n}^{\infty} y_i \psi_i.$$

If S is the unit sphere in Φ, then for every $\epsilon > 0$ there is an n_ϵ such that for all $n > n_\epsilon$ and $x \epsilon S$,

$$\| y^{(n)} \| < \frac{\epsilon}{2}$$

because of Theorem 2. For any fixed $n > n_\epsilon$ there is an $x(n) \epsilon S$ such that

$$| T_n | \leqq \| T_n x(n) \| + \frac{\epsilon}{2},$$

so that

$$| T_n | \leqq \| y^{(n)}(n) \| + \frac{\epsilon}{2} < \epsilon,$$

where $y(n) = Tx(n)$.

With a view to obtaining results on absolutely convergent determinants we state certain sufficient conditions in terms of the rows and the columns of the matrix $(a_{ij}) = (T_i^{\Psi} T\varphi_j)$ separately and impose the following conditions on Φ.

(a) If $x \epsilon \Phi$, then $x^+ = \sum_{n=1}^{\infty} | T_n^{\Phi} x | \varphi_n \epsilon \Phi$ and $\| x^+ \| = \| x \|$.

(b) If $0 \leqq T_n^{\Phi} x \leqq T_n^{\Phi} x'$, then $\| x \| \leqq \| x' \|$.

These conditions are satisfied by the spaces l_p and c_0.

THEOREM 6. *If T is linear on Φ to Ψ and*

$$L = \sum_{i=1}^{\infty} | L_i | \psi_i \epsilon \Psi,$$

then T is completely continuous.

Proof. Let T_n be defined by the matrix

$$\begin{aligned} a_{ij} &= T_i^{\Psi} T\varphi_j, && 1 \leqq i \leqq n, \\ &= 0, && i > n. \end{aligned}$$

Since $T_n x$ lies in a finite-dimensional subspace of Ψ, T_n is completely continuous. For any $\epsilon > 0$,

$$\begin{aligned} \| Tx - T_n x \| &= \left\| \sum_{i=n+1}^{\infty} y_i \psi_i \right\| = \left\| \sum_{i=n+1}^{\infty} | y_i | \psi_i \right\| \\ &\leqq \| x \| \cdot \left\| \sum_{i=n+1}^{\infty} | L_i | \psi_i \right\| < \epsilon \| x \|, \end{aligned}$$

if $n > n_\epsilon$. Thus T is completely continuous by Theorem 4.

COROLLARY. *If* $\sum_{i=1}^{\infty} \left(\sum_{j=1}^{\infty} | a_{ij} |^{p'} \right)^{q/p'} < \infty$, *then* $y_i = \sum_{j=1}^{\infty} a_{ij} x_j$ *is completely continuous on l_p to l_q ($p + p' = pp'$).*

THEOREM 7. *If T is linear on Φ to Ψ and*

$$\sum_{j=1}^{\infty} \| \psi_j^* \| x_j$$

is a linear functional on Φ such that for $\epsilon > 0$

$$\sum_{j=n+1}^{\infty} \| \psi_j^* \| \cdot | x_j | < \epsilon \| x \| \qquad (n < n_\epsilon),$$

then T is completely continuous.

Proof. Let T_n be defined by the matrix

$$\begin{aligned} a_{ij} &= T_i^{\Psi} T\varphi_j, && 1 \leqq j \leqq n, \\ &= 0, && j > n. \end{aligned}$$

T_n, defined on an n-dimensional space, is completely continuous. Now for $\epsilon > 0$, $n > n_\epsilon$ and $k > 0$

$$\left\| \sum_{j=n+1}^{n+k} x_j \psi_j^* \right\| \leqq \sum_{j=n+1}^{n+k} | x_j | \cdot \| \psi_j^* \| \leqq \sum_{j=n+1}^{\infty} | x_j | \cdot \| \psi_j^* \| < \epsilon \| x \|.$$

Further, $T_n x = Tx^{(n)} = T \sum_{j=1}^{n} x_j \varphi_j = \sum_{j=1}^{n} x_j \psi_j^*$, so that

$$\| Tx - T_n x \| = \left\| \sum_{j=n+1}^{\infty} x_j \psi_j^* \right\| < \epsilon \| x \|.$$

COROLLARY. *If* $\sum_{j=1}^{\infty}\left(\sum_{i=1}^{\infty}|a_{ij}|^q\right)^{p'/q} < \infty$, *then* $y_i = \sum_{j=1}^{\infty} a_{ij}x_j$ *is completely continuous on* l_p *to* l_q.

The next stage of restriction on T is made to assure the absolute convergence of the determinant of the matrix of $I + T$. The condition is essentially the combination of the conditions of the last two theorems.

THEOREM 8. *The determinant of the matrix* $(\delta_{ij} + a_{ij})$ *defined by* $I + T$ *on* Φ *to* Φ *is absolutely convergent if*

1. $L_i x = \sum_{j=1}^{\infty} |a_{ij}| x_j$ *is a linear functional on* Φ;

2. $L = \sum_{i=1}^{\infty} |L_i| \varphi_i \in \Phi$, $\|L\| < 1$;

3. $\varphi x = \sum_{j=1}^{\infty} \|\varphi_j^*\| x_j$ *is a linear functional on* Φ, $\varphi_j^* = T\varphi_j$;

4. $\sum_{i=1}^{\infty} |a_{ii}| = a < \infty$.

Proof. According to the theorem of von Koch it is enough to show that

$$\sum_{i=1}^{\infty}|a_{ii}| + \sum_{n=1}^{\infty}\sum_{ii_1\cdots i_n=1}^{\infty}|a_{ii_1}a_{i_1i_2}\cdots a_{i_ni}|$$

converges. We show first that for all i, j, n

$$\sum_{i_1\cdots i_n=1}^{\infty}|a_{ii_1}a_{i_1i_2}\cdots a_{i_nj}| \leqq |L_i|\cdot\|L\|^{n-1}\cdot\|\varphi_j^*\|. \tag{1}$$

For $n = 1$,

$$\sum_{i_1=1}^{\infty}|a_{ii_1}a_{i_1j}| = L_i\varphi_j^{*+} \leqq |L_i|\cdot\|\varphi_j^*\|.$$

From the inductive assumption we have

$$\begin{aligned}\sum_{i_1\cdots i_{n+1}=1}^{\infty}|a_{ii_1}a_{i_1i_2}\cdots a_{i_{n+1}j}| &= \sum_{i_1=1}^{\infty}|a_{ii_1}|\sum_{i_2\cdots i_{n+1}=1}^{\infty}|a_{i_1i_2}\cdots a_{i_{n+1}j}|\\ &\leqq \sum_{i_1=1}^{\infty}|a_{ii_1}|\cdot|L_{i_1}|\cdot\|L\|^{n-1}\cdot\|\varphi_j^*\| = \|L\|^{n-1}\cdot\|\varphi_j^*\|\,L_iL\\ &\leqq |L_i|\cdot\|L\|^n\cdot\|\varphi_j^*\|.\end{aligned}$$

Putting $i = j$ in (1) we have

$$\begin{aligned}\sum_{i=1}^{\infty}\sum_{i_1\cdots i_n=1}^{\infty}|a_{ii_1}a_{i_1i_2}\cdots a_{i_ni}| &\leqq \|L\|^{n-1}\sum_{i=1}^{\infty}\|\varphi_i^*\|\cdot|L_i|\\ &= \|L\|^{n-1}\varphi L \leqq |\varphi|\cdot\|L\|^n.\end{aligned}$$

Thus

$$\sum_{i=1}^{\infty} |a_{ii}| + \sum_{n=1}^{\infty} \sum_{i i_1 \cdots i_n = 1}^{\infty} |a_{ii_1} a_{i_1 i_2} \cdots a_{i_n i}| \leqq a + |\varphi| \sum_{n=1}^{\infty} ||L||^n = a + \frac{|\varphi| \cdot ||L||}{1 - ||L||},$$

and the theorem is proved.

The matrix of minors A_{ij} has certain properties similar to those of the matrix a_{ij}.

THEOREM 9. $\sum_{j=1}^{\infty} | A_{ij} | \varphi_j \epsilon \Phi$ *for each* i.

$\sum_{i=1}^{\infty} | A_{ij} | x_i$ *is a linear functional on* Φ *for each* j.

Proof. Let A_{rc} be the minor of a_{rc} $(r \neq c)$. The matrix of A_{rc}, after a finite number of interchanges of rows and of columns has the form

$$\begin{array}{c|c} a_{cr} & (a_{ir}) \ (i \neq c) \\ \hline (a_{cj}) & (\delta_{ij} + a_{ij}) \\ (j \neq r) & (i \neq c, \ j \neq r) \end{array}.$$

If $P = \Pi(1 + | p |)$ is the product extended over all circular products

$$p = a_{ii_1} a_{i_1 i_2} \cdots a_{i_n i} \qquad (i, i_1, \cdots, i_n \text{ distinct}),$$

then

$$| A_{rc} | \leqq P\{| a_{cr} | + \sum_{n=1}^{\infty} \sum_{i_1 \cdots i_n = 1}^{\infty} | a_{ci_1} a_{i_1 i_2} \cdots a_{i_n r} |\}.$$

From (1) of the previous proof

$$| A_{rc} | \leqq P\left\{| a_{cr} | + \frac{| L_c | \cdot || \varphi_r^* ||}{1 - || L ||}\right\}.$$

Since $| a_{rc} | \leqq | L_c |, || \varphi_r^* ||$, we have

$$| A_{rc} | \leqq | L_c | P\left\{1 + \frac{|| \varphi_r^* ||}{1 - || L ||}\right\},$$

$$| A_{rc} | \leqq || \varphi_r^* || P\left\{1 + \frac{| L_c |}{1 - || L ||}\right\}$$

for $r \neq c$. But $| A_{ii} | \leqq P$ for all i. Hence under the postulates (a) and (b) the theorem holds.

We turn to a consideration of the spaces l_p and c_0 in order to determine the exact conditions for continuity and complete continuity of the transformation and to evaluate the norm in certain cases.

THEOREM 10. *The equations*

(1) $$y_i = \sum_{j=1}^{\infty} a_{ij} x_j$$

define a linear transformation on l_1 *to* l_p *if and only if*

(a) $$\sup_j \left(\sum_{i=1}^{\infty} | a_{ij} |^p \right)^{1/p} < \infty,$$

and this constant is the norm of the transformation.

Proof. Let φ_j be the vector with one in the j-th place and zero elsewhere. Then

$$\sup_j \left(\sum_{i=1}^{\infty} | a_{ij} |^p \right)^{1/p} = \sup_{\|x\|=1} \sum_{j=1}^{\infty} | x_j | \left(\sum_{i=1}^{\infty} | a_{ij} |^p \right)^{1/p}$$
$$\geqq \sup_{\|x\|=1} \left(\sum_{i=1}^{\infty} \left| \sum_{j=1}^{\infty} a_{ij} x_j \right|^p \right)^{1/p} = | T |$$
$$\geqq \sup_j \| T\varphi_j \| = \sup_j \left(\sum_{i=1}^{\infty} | a_{ij} |^p \right)^{1/p}.$$

THEOREM 11. *The equations* (1) *define a linear transformation on* l_1 *to* c_0 *if and only if*

(a′) $\sup_j \sup_i | a_{ij} | < \infty,$ (a″) $\lim_i a_{ij} = 0,$

and the constant (a′) *is the norm of the transformation.*

Proof. Let φ_j be the unit vector with one in the j-th place. Then $\lim_i a_{ij} = 0$ since $T\varphi_j = (a_{ij}) \epsilon c_0$ for each j and

$$\sup_j \sup_i | a_{ij} | = \sup_j \| T\varphi_j \| \leqq | T | = \sup_{\|x\|=1} \sup_i \left| \sum_{j=1}^{\infty} a_{ij} x_j \right| \leqq \sup_i \sup_j | a_{ij} |.$$

THEOREM 12. *The equations* (1) *define a completely continuous transformation on* l_1 *to* l_p $(1 \leqq p < \infty)$ *if and only if the matrix* (a_{ij}) *satisfies the conditions* (a) *and*

(b) $$\lim_n \sup_j \left(\sum_{i=n}^{\infty} | a_{ij} |^p \right)^{1/p} = 0.$$

Proof. The necessity of (a) follows from Theorem 10 and that of (b) from Theorem 3. The conditions are sufficient since

$$\left(\sum_{i=n}^{\infty} | y_i |^p \right)^{1/p} \leqq \sup_j \left(\sum_{i=n}^{\infty} | a_{ij} |^p \right)^{1/p} \sum_{i=1}^{\infty} | x_i |$$

and Theorem 2 show that bounded sets go into compact sets.

THEOREM 13. *The equations* (1) *define a completely continuous transformation on* l_1 *to* c_0 *if and only if the matrix* (a_{ij}) *satisfies the conditions* (a′) *and*

(b′) $\lim_i a_{ij} = 0$ *uniformly in* j.

Proof. The necessity of these conditions follows from Theorems 11 and 3. The conditions are sufficient since

$$\sup_{i \geqq n} |y_i| \leqq \sup_{i \geqq n} \sup_j |a_{ij}| \sum_{i=1}^{\infty} |x_i|$$

and Theorem 2 show that the image of a bounded set is compact.

The rôles of l_1 and l_p or c_0 in these theorems may be interchanged if the a_{ij} are positive and the norms of the transformations evaluated. The evaluations are corollaries of the following theorem in which the postulates (a) and (b) on Φ used in Theorems 8 and 9 appear in a weaker form as one restriction on Φ and one on the conjugate space $\bar{\Phi}$.

THEOREM 14. *Let $a_{ij} \geqq 0$ and Φ be a space with a basis φ_n such that*

(i) *if $\sum_i x_i\varphi_i$ converges so does $\sum_i |x_i| \varphi_i$;*

(ii) *the norm of a point*[3] *$(\alpha_j) \in \bar{\Phi}$ is not decreased when the values $|\alpha_j|$ increase. Then equations* (1) *define a linear transformation on Φ to l_1 if and only if $\sum_{i=1}^{\infty} a_{ij} \in \bar{\Phi}$ (i.e., if and only if $\sum_{j=1}^{\infty}\sum_{i=1}^{\infty} a_{ij}x_j$ converges whenever $\sum_j x_j\varphi_j$ converges). Furthermore the constant*

$$\left\| \sum_{i=1}^{\infty} a_{ij} \right\|_{\bar{\Phi}}$$

is the norm of the transformation.

Proof. Let $f(y)$ be the linear functional on l_1 defined by

$$f(y) = \sum_{i=1}^{\infty} y_i.$$

If equations (1) define a linear transformation $y = Tx$ on Φ to l_1, then fT is a linear functional on Φ with $|fT| \leqq |T|$. But

$$fTx = \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} a_{ij}x_j = \sum_{j=1}^{\infty} \alpha_j x_j.$$

If (x_j) is the unit vector with one in the j-th place, we see that

$$(\alpha_j) = \left(\sum_{i=1}^{\infty} a_{ij}\right) \in \bar{\Phi}$$

and

$$\left\| \sum_{i=1}^{\infty} a_{ij} \right\|_{\bar{\Phi}} \leqq |T|.$$

[3] If $x = \sum_j x_j\varphi_j$ and f is in $\bar{\Phi}$, then $fx = \sum_j x_j\alpha_j$, where $\alpha_j = f\varphi_j$.

In what follows f will stand for an arbitrary linear functional on l_1. From a well-known theorem (Banach, *Théorie des Opérations Linéaires*, page 55, Theorem 3) we have

$$\sup_{|f|=1} |fT| = |T|.$$

Now f is represented by a vector (f_i) in the space of bounded sequences and $|f| = \sup |f_i|$. Also

$$fTx = f\sum_{j=1}^{\infty} a_{ij}x_j = \sum_{j=1}^{\infty} f(a_{ij})x_j = \sum_{j=1}^{\infty}\sum_{i=1}^{\infty} f_i a_{ij} x_j,$$

so that

$$|fT| = \left\| \sum_{i=1}^{\infty} f_i a_{ij} \right\|_{\bar{\Phi}}.$$

Thus

$$|T| = \sup_{|f|=1} \left\| \sum_{i=1}^{\infty} f_i a_{ij} \right\|_{\bar{\Phi}} = \sup_{|f_i| \leqq 1} \left\| \sum_{i=1}^{\infty} f_i a_{ij} \right\|_{\bar{\Phi}} \leqq \left\| \sum_{i=1}^{\infty} a_{ij} \right\|_{\bar{\Phi}}.$$

Hence we have shown that if $y = Tx$ is linear on Φ to l_1, then $\sum_{i=1}^{\infty} a_{ij} \in \bar{\Phi}$ and $|T| = \left\| \sum_{i=1}^{\infty} a_{ij} \right\|_{\bar{\Phi}}$. Now conversely if $\sum_{i=1}^{\infty} a_{ij} \in \bar{\Phi}$, we have the iterated sum

$$\sum_{j=1}^{\infty}\sum_{i=1}^{\infty} a_{ij}x_j$$

converging absolutely and hence for every $x \in \Phi$ the sequence

$$\sum_{j=1}^{\infty} a_{ij} T_j^{\Phi} x \in l_1.$$

By Theorem 1 then the transformation is continuous.

COROLLARY 1. *In case the $a_{ij} \geqq 0$, the equations* (1) *define a linear transformation on l_p $(p > 1)$ to l_1 if and only if*

$$\left(\sum_{j=1}^{\infty} \left(\sum_{i=1}^{\infty} a_{ij} \right)^{p'} \right)^{1/p'} < \infty.$$

This constant is the norm of the transformation.

COROLLARY 2. *In case $a_{ij} \geqq 0$, the equations* (1) *define a linear transformation on c_0 to l_1 if and only if*

$$\sum_{j=1}^{\infty}\sum_{i=1}^{\infty} a_{ij} < \infty.$$

This constant is the norm of the transformation.

Using these corollaries, the corollary to Theorem 7, and Theorem 2 we get

COROLLARY 3. *A linear transformation with $a_{ij} \geqq 0$ on l_p $(1 < p < \infty)$ or c_0 to l_1 is necessarily completely continuous.*

Part of this corollary but not the evaluation of the norm given in Corollary 1 has been proved in more general form by H. R. Pitt.[4] He has shown that any linear transformation on l_p to l_q is necessarily completely continuous if $p > q$. Littlewood[5] has given the same result for $p = \infty$, $q = 1$.

COROLLARY 4. *If T is linear on c_0 to l_1, and $a_{ij} \geqq 0$, then the determinant of $I + T$ is absolutely convergent.*

In fact, it follows from Corollary 2 that the determinant of $(\delta_{ij} + a_{ij})$ is a normal determinant in the sense of von Koch.

We conclude with the representation of the general completely continuous operation on L (the space of functions $\varphi(P)$ summable on $(0, 1)$) to l_q $(1 \leqq q < \infty)$.

THEOREM 15. *The function T is a completely continuous linear operation on L to l_p if and only if it is expressible in the form*

$$T\varphi = \int_0^1 k_i(P)\varphi(P)\,dP \tag{a}$$

with measurable functions $k_i(P)$ satisfying

$$\operatorname*{ess\,sup}_P \left(\sum_{i=1}^{\infty} | k_i(P) |^q\right)^{1/q} < \infty, \tag{b}$$

$$\lim_n \operatorname*{ess\,sup}_P \left(\sum_{i=n}^{\infty} | k_i(P) |^q\right) = 0. \tag{c}$$

The constant in (b) *is the norm of the transformation.*

Proof. It is known[6] that $T\varphi$ is linear on L to l_q if and only if (a) and (b) are satisfied and that (b) gives the norm of the transformation. Let $T_n\varphi$ be defined as the vector

$$\left(0, 0, \cdots, \int_0^1 k_n(P)\varphi(P)\,dP, \int_0^1 k_{n+1}(P)\varphi(P)\,dP, \cdots\right).$$

Then

$$\| T_n\varphi \| \leqq \operatorname*{ess\,sup}_P \left(\sum_{i=n}^{\infty} | k_i(P) |^q\right)^{1/q} \| \varphi \|.$$

Corollary 1 of Theorem 2 then shows that (c) is sufficient for the complete continuity of T. Conversely, if T is completely continuous, we have for every $\epsilon > 0$ an $n(\epsilon)$ such that for every $\varphi \in L$ with $\| \varphi \| \leqq 1$

$$\| T_n\varphi \| = \left\{\sum_{i=n}^{\infty} \left| \int_0^1 k_i(P)\varphi(P)\,dP \right|^q\right\}^{1/q} \leqq \epsilon \qquad (n \geqq n(\epsilon)).$$

[4] H. R. Pitt, *A note on bilinear forms*, Journal of the London Math. Soc., vol. 11 (1936), pp. 174–180.

[5] Loc. cit.

[6] See Dunford, *Integration and linear operations*, Trans. Amer. Math. Soc., vol. 40 (1936), p. 486, or B. Vullich, *Sur les opérations linéaires dans l'espace des fonctions sommables*, Mathematica, vol. 13 (1937), p. 42, Theorem 1.

There exists a φ_n in L with $\| \varphi_n \| = 1$ such that

$$| T_n | \leqq \| T_n \varphi_n \| + \epsilon.$$

Thus for $n \geqq n(\epsilon)$

$$| T_n | = \operatorname*{ess\,sup}_P \left\{ \sum_{i=n}^{\infty} | k_i(P) |^q \right\}^{1/q} \leqq 2\epsilon,$$

and so (c) is also necessary.

UNIVERSITY OF KENTUCKY AND YALE UNIVERSITY.

ON PERFECT METHODS OF SUMMABILITY

By J. D. Hill

1. **Introduction.** In this paper we are concerned exclusively with Toeplitz methods of summability in the real domain, and we begin by introducing the definitions and notations which we shall employ. Being given a matrix $A = (a_{nk})$ $(k, n = 0, 1, 2, \cdots)$ and a sequence $x = \{s_k\}$, we may form the new sequence $y \equiv A(x) \equiv \{t_n\}$ provided each of the series $\sum_{k=0}^{\infty} a_{nk} s_k \equiv t_n \equiv A_n(x)$ is convergent. If y belongs to the space (c) of convergent sequences, we say that x is *summable by the method* A, or simply A-*summable*, and we write A-lim $x =$ lim y. The class $[A]$ of all A-summable sequences is called the *convergence-field* of A. If for two methods A and B we have the relation $[A] \subset [B]$, we say that B *is not weaker than* A. A and B are said to be *consistent* if A-lim $x = B$-lim x whenever these limits exist. The method I defined by the matrix (δ_{nk}), where δ_{nk} is Kronecker's symbol, is called the *identical method* or the *identity*; obviously $[I] = (c)$. Every method A for which $[I] \subset [A]$ is called *convergence-preserving*; if, in addition, A is consistent with I, it is said to be *regular*. If the matrix (a_{nk}) is such that $a_{nk} = 0$ for $k > n$, A is said to be *triangular*; if, furthermore, $a_{nn} \neq 0$ for every n, A is said to be *normal*. A will be called *reversible* if the equation $A(x) = y$ has exactly one solution x, convergent or not, for each y in (c). For triangular methods the notions of reversibility and normality are easily seen to be equivalent.

For future reference we list here the following conditions which are necessary and sufficient for A to be regular:

$$\lim_{n\to\infty} a_{nk} = 0 \qquad (k = 0, 1, 2, \cdots), \tag{1.1}$$

$$\lim_{n\to\infty} \sum_{k=0}^{\infty} a_{nk} = 1, \tag{1.2}$$

$$\sum_{k=0}^{\infty} |a_{nk}| \leq K \qquad (n = 0, 1, 2, \cdots). \tag{1.3}$$

We shall say[1] that A is of *type M* if the conditions

$$\sum_{n=0}^{\infty} |\alpha_n| < \infty, \qquad \sum_{n=0}^{\infty} \alpha_n a_{nk} = 0 \qquad (k = 0, 1, 2, \cdots) \tag{1.4}$$

Received August 16, 1937. The author gratefully acknowledges his indebtedness to Professor J. D. Tamarkin, under whose direction this paper was written.

[1] Matrices of this type were first introduced by Mazur in connection with normal methods; see *Eine Anwendung der Theorie der Operationen bei der Untersuchung der Toeplitzschen Limitierungsverfahren*, Studia Mathematica, vol. 2 (1930), pp. 40–50. We shall refer to this paper hereafter as SM.

always imply

$$\alpha_n = 0 \qquad (n = 0, 1, 2, \cdots). \tag{1.5}$$

Banach[2] calls a method *perfect* if it is simultaneously regular, reversible, and of type M. The importance of perfect methods lies in the following theorems.

THEOREM 1 (Mazur).[3] *In order that a normal regular method A be consistent with every regular method not weaker than A, it is necessary and sufficient that A be of type M.*

Banach has shown that the sufficiency may be extended to methods which are not necessarily normal.

THEOREM 2 (Banach).[4] *In order that A be consistent with every regular method not weaker than A, it is sufficient that A be perfect.*

The only known examples of perfect methods are the Cesàro and Euler methods of all positive orders, a result obtained by Mazur in SM. It is the purpose of the present paper to find conditions under which the Nörlund, Hausdorff, and weighted-mean methods will be of type M. We start, however, with a few theorems of a general nature.

2. Perfect methods in general. Let $A = (a_{nk})$ be a given regular matrix. If we regard $A(x)$ as an operation on (c), the regularity insures that its range R_c will be a subset of (c), and we then have the following characterization of perfect methods.

THEOREM 3. *In order that a regular and reversible method A be of type M, it is necessary and sufficient that R_c be dense in (c).*

Proof. The assertion of the necessity is merely an alternative statement of a result due to Banach.[5] To establish the sufficiency consider the points $y_i \equiv \{\delta_{ni}\}$ of (c), $(i = 0, 1, 2, \cdots)$. If R_c is dense in (c), there exists corresponding to each $\epsilon > 0$ and each $i = 0, 1, 2, \cdots$ a convergent (and hence bounded) sequence $\{s_{ki}\} \equiv x_i$, such that $\| A(x_i) - y_i \| < \epsilon$, which is equivalent to

$$| A_n(x_i) - \delta_{ni} | < \epsilon \qquad (n, i = 0, 1, 2, \cdots). \tag{2.1}$$

We write (2.1) in the form

$$A_n(x_i) = \delta_{ni} + \epsilon_{ni}, \text{ where } | \epsilon_{ni} | < \epsilon, \tag{2.2}$$

and assume that (1.4) holds. Then for each fixed i it is clear from (1.3) and the boundedness in k of $\{s_{ki}\}$ that $\sum_{k,n=0}^{\infty} \alpha_n a_{nk} s_{ki}$ is absolutely convergent. Conse-

[2] *Théorie des Opérations Linéaires*, p. 90.

[3] See SM, p. 48, Satz 7. It may be remarked here that Mazur (see Mathematische Zeitschrift, vol. 28 (1928), pp. 604–605) has constructed two normal regular methods, each not weaker than the other, which assign different limits to a particular sequence.

[4] Loc. cit., p. 95, Théorème 12.

[5] Loc. cit., p. 93, Lemme 2.

quently, from (1.4) and (2.2) we have

$$0 = \sum_{k=0}^{\infty} \sum_{n=0}^{\infty} \alpha_n a_{nk} s_{ki} = \sum_{n=0}^{\infty} \left\{ \sum_{k=0}^{\infty} a_{nk} s_{ki} \right\} \alpha_n = \sum_{n=0}^{\infty} A_n(x_i) \alpha_n$$

$$= \sum_{n=0}^{\infty} \delta_{ni} \alpha_n + \sum_{n=0}^{\infty} \epsilon_{ni} \alpha_n = \alpha_i + \sum_{n=0}^{\infty} \epsilon_{ni} \alpha_n .$$

Therefore $|\alpha_i| \leqq \epsilon \cdot \sum_{n=0}^{\infty} |\alpha_n|$, and, $\epsilon > 0$ being arbitrary, it follows that $\alpha_i = 0$ $(i = 0, 1, 2, \cdots)$. This completes the proof.

On the other hand, if we consider $A(x)$ as an operation defined in the space (m) of bounded sequences, its range R_m lies in (m) on account of (1.3), and it is readily seen from the foregoing proof that the following theorem holds.

THEOREM 4. *In order that a regular and reversible method A be of type M it is necessary and sufficient that the points $y_i \equiv \{\delta_{ni}\}$ of (m) be limit points of R_m for all $i = 0, 1, 2, \cdots$.*

THEOREM 5. *The product $AB \equiv C$ of two triangular perfect methods A and B is also a triangular perfect method.*

Proof. If the matrices of A, B, and C are, respectively, (a_{nk}), (b_{nk}), and (c_{nk}), then $c_{nk} = \sum_{i=k}^{n} a_{ni} b_{ik}$ if $k \leqq n$, and $c_{nk} = 0$ if $k > n$, so that C is triangular. Moreover, it is obvious that C is normal and regular. Conditions (1.4) applied to C give

$$\sum_{n=k}^{\infty} \alpha_n c_{nk} = \sum_{n=k}^{\infty} \sum_{i=k}^{n} \alpha_n a_{ni} b_{ik} = \sum_{i=k}^{\infty} \left(\sum_{n=i}^{\infty} \alpha_n a_{ni} \right) b_{ik} = 0 \quad (k = 0, 1, 2, \cdots), \tag{2.3}$$

where the interchange of summation signs is permitted by the absolute convergence of $\sum_{i,n=0}^{\infty} \alpha_n a_{ni} b_{ik}$. Setting

$$\beta_i \equiv \sum_{n=i}^{\infty} \alpha_n a_{ni} \qquad (i = 0, 1, 2, \cdots), \tag{2.4}$$

we have $\sum_{i=0}^{\infty} |\beta_i| < \infty$ and $\sum_{i=k}^{\infty} \beta_i b_{ik} = 0$ for $k = 0, 1, 2, \cdots$. Since B is of type M, this implies $\beta_i = 0$ for $i = 0, 1, 2, \cdots$. This in turn, by (2.4) and the fact that A is of type M, implies $\alpha_n = 0$ for every n. Thus C is also of type M, and the proof is complete.

THEOREM 6. *If the product $AB \equiv C$ of two triangular convergence-preserving methods A and B is of type M, then A must be of type M.*

Proof. If A is not of type M, there must exist a sequence $\{\alpha_n'\}$ satisfying (1.4) but not (1.5). Since (1.3) is also necessary for preservation of convergence, the relation (2.3) can then be established for the sequence $\{\alpha_n'\}$. This contradicts the assumption that C is of type M.

We shall have occasion later to use the following theorem, the proof of which is easily supplied.

THEOREM 7. *If A is normal and B is triangular, then B is not weaker than A if and only if BA^{-1} is convergence-preserving.*

3. **The Nörlund method.** The Nörlund method of summation corresponds to a triangular matrix whose elements are of the form $a_{nk} \equiv p_{n-k}/P_n$ for $k = 0, 1, \cdots, n$; $n = 0, 1, 2, \cdots$, where $\{p_k\}$ is a given sequence such that $P_n \equiv p_0 + p_1 + \cdots + p_n \neq 0$ for all n. In particular, for $n = 0$ we have $p_0 \neq 0$, and we may always take $p_0 = 1$ since p_k may be replaced by p_k/p_0 without affecting the matrix. Then in view of $a_{nn} \equiv 1/P_n \neq 0$ we see that every Nörlund matrix is normal. Furthermore, since $\sum_{k=0}^{n} p_{n-k}/P_n = 1$ $(n = 0, 1, 2, \cdots)$, the regularity conditions reduce to (1.1) and (1.3) alone. In the special case where $p_k \geqq 0$ for every k, (1.3) is always fulfilled and on account of $p_{n-k}/P_n \leqq p_{n-k}/P_{n-k}$ the conditions become simply $\lim_{n\to\infty} p_n/P_n = 0$.

The question naturally arises whether a regular Nörlund matrix is necessarily of type M, and the following example shows that this is not the case.

Example 1. Let $p_1 = 2$ and $p_k = 0$ for $k \geqq 2$; since $p_k \geqq 0$ and $p_k/P_k = 0$ if $k \geqq 2$, the corresponding method is regular by the above remark. On the other hand, conditions (1.4), but not (1.5), are satisfied if we let $\alpha_0 = \frac{1}{3}$ and $\alpha_n = (-\frac{1}{2})^n$ for $n \geqq 1$.

THEOREM 8. *For a regular Nörlund method to be of type M it is sufficient that the sequence*

$$D_n \equiv \begin{vmatrix} p_1 & 1 & 0 & 0 & \cdots & 0 \\ p_2 & p_1 & 1 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ p_{n-1} & p_{n-2} & \cdot & \cdot & \cdots & 1 \\ p_n & p_{n-1} & \cdot & \cdot & \cdots & p_1 \end{vmatrix} \qquad (n = 1, 2, 3, \cdots)$$

be bounded.

Proof. We obtain this result as a consequence of Theorem 4 by imposing the condition that R_m actually contain the points y_i. This is equivalent to requiring that the solution $\{s_{ki}\}$ of the system of equations

$$A_n(x_i) \equiv \sum_{k=0}^{n} (p_{n-k}/P_n)s_{ki} = \delta_{ni} \qquad (n = 0, 1, 2, \cdots) \tag{3.1}$$

be bounded in k for each fixed $i = 0, 1, 2, \cdots$. Observing that the first i of the s_{ki} are zero, we may write (3.1) in the form

$$\sum_{k=0}^{n} p_{n-k}s_{i+k,i} = P_{i+n}\delta_{i+n,i} \qquad (n = 0, 1, 2, \cdots) \tag{3.2}$$

from which we obtain by applying Cramer's rule to the first $j + 1$ of these equations

$$s_{i+j,i} = (-1)^j P_i D_j \qquad (i, j = 0, 1, 2, \cdots; D_0 = 1), \tag{3.3}$$

and the theorem follows.

By condition (1.1) the regularity of a Nörlund method implies that $p_{n+1}/P_{n+1} = (P_{n+1} - P_n)/P_{n+1} = 1 - (P_n/P_{n+1}) \to 0$, or that $P_n/P_{n+1} \to 1$

as $n \to \infty$. Thus the power series $\sum_{n=0}^{\infty} P_n z^n$ has a radius of convergence equal to 1, and consequently the radius of convergence of $p(z) \equiv \sum_{n=0}^{\infty} p_n z^n \equiv (1 - z) \sum_{n=0}^{\infty} P_n z^n$ is at least 1. Combining (3.2) and (3.3) gives the relation $\sum_{k=0}^{n} (-1)^k p_{n-k} D_k = \delta_{n0}$ for $n = 0, 1, 2, \cdots$, from which we obtain the formula

$$1/p(z) = \sum_{n=0}^{\infty} (-1)^n D_n z^n, \tag{3.4}$$

valid for $|z|$ sufficiently small since $p(0) = 1$. Hence for each regular Nörlund method there exist positive constants A and B such that $|D_n| \leqq AB^n$ for all n, and the condition of Theorem 8 will be satisfied if and only if B can be chosen $\leqq 1$. That the latter is by no means a necessary condition is shown by the following example.

Example 2. For each $r = 1, 2, 3, \cdots$ there exists a perfect Nörlund method for which the corresponding D_n is precisely $O(n^{r-1})$. For let r be chosen arbitrarily and fixed. Define p_n as $\binom{r}{n}$ for $n = 0, 1, \cdots, r$ and as zero otherwise. We then have $p(z) = (1 + z)^r$ and $1/p(z) = \sum_{n=0}^{\infty} (-1)^n \binom{n + r - 1}{r - 1} z^n$, so that by (3.4), $D_n = \binom{n + r - 1}{r - 1} = O(n^{r-1})$.

The p_n being non-negative and ultimately zero, the regularity is apparent. Let us then assume that conditions (1.4) are satisfied. For $k \geqq r$ these reduce to

$$L_r(\alpha_k) \equiv \sum_{i=0}^{r} \binom{r}{i} \alpha_{k+i} = 0 \qquad (k = r, r + 1, r + 2, \cdots). \tag{3.5}$$

The latter implies $L_{r-1}(\alpha_k) = 0$ for all $k \geqq r$. For we have evidently

$$L_r(\alpha_k) = L_{r-1}(\alpha_k) + L_{r-1}(\alpha_{k+1}) = 0,$$

$$L_r(\alpha_{k+1}) = L_{r-1}(\alpha_{k+1}) + L_{r-1}(\alpha_{k+2}) = 0,$$

whence $L_{r-1}(\alpha_k) = L_{r-1}(\alpha_{k+2})$. Since $\alpha_n \to 0$ it is clear that $\{L_{r-1}(\alpha_{k+2i})\}$ for each fixed $k \geqq r$ is a sequence of equal terms converging to zero. This establishes the assertion.

It follows then by induction that conditions (3.5) imply $L_0(\alpha_k) \equiv \alpha_k = 0$ for all $k \geqq r$. The first r equations in (1.4) now take a simplified form from which it is obvious that $\alpha_k = 0$ for $k = 0, 1, \cdots, r - 1$, and the example is complete.

We conclude this section with a few instances in which the criterion of Theorem 8 applies.

Mazur's result for Cesàro summability $(C, \alpha > 0)$ is immediate. For we have

$p_n = \binom{n + \alpha - 1}{n}$, $p(z) = (1 - z)^{-\alpha}$, and $1/p(z) = \sum_{n=0}^{\infty} (-1)^n \binom{\alpha}{n} z^n$; thus $D_n = \binom{\alpha}{n} \to 0$ as $n \to \infty$.

If $p_n = 1/(n + 1)$, we obtain the so-called logarithmic Nörlund method. In this case $1/p(z) = -z/\log(1 - z) \equiv \sum_{n=0}^{\infty} (-1)^n D_n z^n$, and it is known that the coefficients D_n in this expansion are[6] $O(1/(n \log^2 n))$.

Finally, consider the method defined by the sequence $p_n = 1 + nd$ for a given $d > 0$. Since $P_n = (n + 1)(nd + 2)/2$, the regularity follows at once. Moreover, this method is definitely stronger than the identity since one may show that it assigns the limit $\frac{1}{2}$ to the divergent sequence $\{(1 + (-1)^n)/2\}$. We have here $p(z) = (1 + (d - 1)z)/(1 - z)^2$ and

$$1/p(z) = 1 - (1 + d)z + d^2 \sum_{n=2}^{\infty} (1 - d)^{n-2} z^n,$$

whence $D_n = d^2(d - 1)^{n-2}$ for $n \geqq 2$. Thus if $0 \leqq d \leqq 2$, this method will be perfect; for other values of d the question remains undecided.

4. **The Hausdorff method.** Let $\{\mu_k\}$ be an arbitrarily given sequence and consider the matrices $S = (s_{nk})$, $T = (t_{nk})$, where $s_{nk} \equiv (-1)^k \binom{n}{k}$, $t_{nk} \equiv \mu_k \delta_{nk}$. Any matrix of the form $H = STS$ is called a Hausdorff matrix. Such a matrix is clearly triangular, and if $H = (h_{nk})$ we find that

$$h_{nk} = \sum_{i=k}^{n} (-1)^{i-k} \binom{n}{k} \binom{n - k}{i - k} \mu_i \qquad (k = 0, 1, \cdots, n; n = 0, 1, 2, \cdots). \tag{4.01}$$

Since[7] $S^2 = I$ it is easy to verify that

(i) *The multiplication of Hausdorff matrices is commutative and the result is again a Hausdorff matrix.*

(ii) *The inverse of a normal Hausdorff matrix is also a (normal) Hausdorff matrix.*

The question whether or not every normal regular Hausdorff matrix is of type M seems to be quite difficult and still remains open. However, we shall later exhibit (see Example 3) a regular Hausdorff matrix which is not of type M. The next theorem, although based on an unverified hypothesis, seems of sufficient interest to warrant its inclusion.

THEOREM 9. *If H_0 is normal, convergence-preserving, and not of type M, and if H is not weaker than H_0, then H is not of type M.*

[6] See Tamarkin, Problem 3276, American Mathematical Monthly, vol. 35 (1928), pp. 497–500; esp. bottom of p. 500.

[7] For a particularly simple proof see Henriksson, *Über die Hausdorffschen Limitierungsverfahren, die schwächer sind als das Abelsche*, Mathematische Zeitschrift, vol. 39 (1935), pp. 501–510; in particular, p. 502.

Proof. By Theorem 7, $H_1 = HH_0^{-1}$ is convergence-preserving and from (i), (ii) we see that $H_1 = H_0^{-1}H$. Hence $H = H_0H_1$ and the result follows at once from Theorem 6.

On the other hand, since Cesàro summability is a special case of the Hausdorff, we know that there exist Hausdorff matrices which are of type M. Concerning such matrices we have the following result.

THEOREM 10. *If H is normal and convergence-preserving, and if H_0 is of type M and not weaker than H, then H is of type M.*

Proof. As in the previous theorem we have $H_1 = H_0H^{-1} = H^{-1}H_0$, where H_1 is convergence-preserving. Then $H_0 = HH_1$ and the conclusion follows again from Theorem 6.

From (i) and Theorem 5 we obtain directly the following theorem.

THEOREM 11. *The product of a finite number of perfect Hausdorff methods is likewise a perfect Hausdorff method.*

We propose next to establish conditions sufficient for a regular Hausdorff matrix to be of type M. In order that a regular method be defined by (4.01), it is necessary and sufficient[8] that μ_i be of the form $\mu_i \equiv \int_0^1 u^i\,dq(u)$ for $i = 0, 1, 2, \cdots$, where $q(u)$ is a function of bounded variation on the interval $U \equiv (0 \leqq u \leqq 1)$ which is continuous at the point $u = 0$ and which satisfies the condition $q(1) - q(0) = 1$. It is understood throughout this section that $q(u)$ will always denote such a function. With this expression for μ_i, (4.01) reduces to $h_{nk} = \binom{n}{k}\int_0^1 u^k(1-u)^{n-k}\,dq(u)$, and conditions (1.4) become

$$(4.02)\qquad \sum_{n=0}^{\infty} |\alpha_n| < \infty,\quad \sum_{n=k}^{\infty}\int_0^1 \alpha_n\binom{n}{k}u^k(1-u)^{n-k}\,dq(u) = 0\quad (k = 0, 1, 2, \cdots).$$

Since

$$(4.03)\qquad \sum_{k=0}^{n}\binom{n}{k}u^k(1-u)^{n-k} \equiv 1\qquad (n = 0, 1, 2, \cdots)$$

and $\int_0^1 dq(u) = q(1) - q(0) = 1$, we observe for future reference that the following relation is implied by (4.02):

$$(4.04)\qquad \sum_{k=0}^{\infty}\sum_{n=k}^{\infty}\alpha_n h_{nk} = \sum_{n=0}^{\infty}\left(\sum_{k=0}^{n} h_{nk}\right)\alpha_n = \sum_{n=0}^{\infty}\alpha_n = 0.$$

We proceed now to reduce (4.02) to a more convenient form. From (4.03) we have $0 \leqq \binom{n}{k}u^k(1-u)^{n-k} \leqq 1$ on U. Consequently, we see that each of the series

$$(4.05)\qquad g_k(u) \equiv \sum_{n=k}^{\infty}\alpha_n\binom{n}{k}u^k(1-u)^{n-k}\qquad (k = 0, 1, 2, \cdots)$$

[8] See Hausdorff, *Summationsmethoden und Momentfolgen*, Mathematische Zeitschrift, vol. 9 (1921), pp. 74–109, 280–299.

converges absolutely and uniformly on U, and that for every k

$$|g_k(u)| \leqq \sum_{n=0}^{\infty} |\alpha_n|, \tag{4.06}$$

$$g_k(u) = (-1)^k u^k g_0^{(k)}(u)/k!. \tag{4.07}$$

Thus $\{g_k(u)\}$ is a uniformly bounded sequence of functions, continuous on U and analytic for $|u-1|<1$. On account of the uniform convergence of (4.05) we may write the second part of (4.02) in the form

$$\int_0^1 g_k(u)\,dq(u) = 0 \qquad (k = 0, 1, 2, \cdots). \tag{4.08}$$

Now for an arbitrarily fixed t on the interval $0 < t \leqq 1$ it follows from (4.06) that the series $G(t, u) \equiv \sum_{k=0}^{\infty} g_k(u)(1-t)^k$ is uniformly convergent on U. Termwise integration is therefore permissible, and from (4.08) we obtain

$$\int_0^1 G(t, u)\,dq(u) = 0. \tag{4.09}$$

But by (4.07), for $0 < u \leqq 1$, we have $G(t, u) \equiv \sum_{k=0}^{\infty} g_k(u)(1-t)^k = \sum_{k=0}^{\infty} g_0^{(k)}(u)(tu-u)^k/k! = g_0(tu)$. By (4.04), however, $g_0(0) \equiv \sum_{n=0}^{\infty} \alpha_n = 0$, and by (4.05), $g_k(0) = 0$ for $k = 1, 2, 3, \cdots$. This shows that the relation $G(t, u) = g_0(tu)$ holds also for $u = 0$. Making this substitution in (4.09) we obtain finally

$$\int_0^1 g_0(tu)\,dq(u) = 0 \qquad (0 \leqq t \leqq 1), \tag{4.10}$$

and the problem is reduced to finding further conditions on $q(u)$ sufficient to insure that (4.10) shall imply $g_0(u) \equiv 0$.

As a first step in this direction we state the following definition.

Condition C. $q(u)$ will be said to satisfy Condition C if there is an index $r \geqq 0$ such that $q^{(r+2)}(u)$ exists and is bounded on U and if $q^{(r+1)}(1) \neq 0$. We shall denote by m the *smallest* index for which this holds. Obviously we then have

$$q^{(i+1)}(1) = 0 \qquad (i = 0, 1, \cdots, m-1). \tag{4.11}$$

If $q(u)$ satisfies this condition, we may form the sequence $q_0(u) \equiv q(u)$, $q_i(u) \equiv uq'_{i-1}(u)$ for $i = 1, 2, \cdots, m$, and one easily sees that

$$q_i(u) = \begin{cases} \sum_{j=1}^{i} C_{ij} u^j q^{(j)}(u) \\ d^i q(u)/dx^i, \text{ when } u = e^x \end{cases} \qquad (i = 1, 2, \cdots, m), \tag{4.12}$$

where the C_{ij} are certain constants.

THEOREM 12. *If $q(u)$ satisfies Condition* C *and if constants $c_0, c_1, \cdots, c_m$ exist such that*

$$\int_0^1 \left| uq_m''(u) + \sum_{i=0}^{m} c_i q_i'(u) \right| du < | q^{(m+1)}(1) |, \tag{4.13}$$

then $q(u)$ defines a Hausdorff matrix of type M.

Proof. From Condition C, (4.11), and the first half of (4.12), it follows that

$$q_i'(1) = \delta_{im} q^{(m+1)}(1), \tag{4.14}$$

$$q_i''(u) = O(1) \text{ on } U, \qquad (i = 0, 1, \cdots, m). \tag{4.15}$$

Setting $s = tu$ for $0 < t \leqq 1$ and writing $dq(u) = q'(u)\, du$, (4.10) becomes $\int_0^t g_0(s)q'(s/t)\, ds = 0$. In view of (4.15) for $i = 0$ we may differentiate this integral and obtain $\int_0^t g_0(s)q''(s/t)(s/t^2)\, ds = q'(1)g_0(t)$. Recalling (4.14) for $i = 0$ and returning to the variable u, we have $\int_0^1 g_0(tu)uq''(u)\, du = 0$. This added to the original expression (4.10) gives $\int_0^1 g_0(tu)q_1'(u)\, du = 0$. This relation is simply (4.10) with $q(u)$ replaced by $q_1(u)$, and (4.15) allows us to repeat this process until we have

$$\int_0^1 g_0(tu)q_i'(u)\, du = 0 \qquad (i = 0, 1, \cdots, m). \tag{4.16}$$

Repeating the process again with $i = m$ and using (4.14) for the index m, we obtain

$$\int_0^1 g_0(tu)uq_m''(u)\, du = q_m'(1)g_0(t) = q^{(m+1)}(1)g_0(t), \tag{4.17}$$

where, by Condition C, $q^{(m+1)}(1) \neq 0$. Multiplying equations (4.16) by the corresponding constants c_i and adding the results to (4.17) give $q^{(m+1)}(1)g_0(t) = \int_0^1 g_0(tu)(uq_m''(u) + \sum_{i=0}^{m} c_i q_i'(u))\, du$ on the interval $0 \leqq t \leqq 1$. Letting $g_0 \equiv \max | g_0(t) |$ on $(0, 1)$, we obtain the inequality

$$g_0 \, | q^{(m+1)}(1) | \leqq g_0 \int_0^1 \left| uq_m''(u) + \sum_{i=0}^{m} c_i q_i'(u) \right| du.$$

Consequently, if (4.13) holds we must have $g_0 = 0$, and the theorem follows.

COROLLARY 1. *If $q(u)$ has a bounded second derivative, and if a constant $c < 1$ exists such that $uq''(u) + cq'(u) \geqq 0$ on U, then the corresponding Hausdorff matrix is of type M.*

Proof. Since $c < 1$ the relation $0 \leqq \int_0^1 | uq''(u) + cq'(u) |\, du =$

$\int_0^1 uq''(u)\,du + c\int_0^1 q'(u)\,du = q'(1) - 1 + c$ gives $q'(1) > 0$, and the hypotheses of Theorem 12 are therefore satisfied with $m = 0$, $c_0 = c$.

From this fact we obtain at once the following result.

COROLLARY 2. *If $q(u)$ has a bounded non-negative second derivative on U, then the corresponding Hausdorff matrix is of type M.*

It seems desirable to show as a final corollary that the criterion of Theorem 12, together with Theorem 10, enables us to reproduce Mazur's result for Cesàro summability.

COROLLARY 3. *The Cesàro matrices of all positive orders are of type M.*

Proof. In this case we have $q(u) = 1 - (1-u)^\alpha$, where $\alpha > 0$. We assume first that α is an integer. Then it is obvious that $q(u)$ satisfies Condition C with $m = \alpha - 1$. Moreover, $q(u) = \sum_{i=1}^{\alpha} (-1)^{i-1}\binom{\alpha}{i}u^i$, and from the second part of (4.12), $q_s(u) = \sum_{i=1}^{\alpha}(-1)^{i-1}i^s\binom{\alpha}{i}u^i$ for $s = 0, 1, \cdots, \alpha - 1$. Consequently, we have

$$uq_m''(u) + \sum_{s=0}^{m} c_s q_s'(u) = \sum_{i=1}^{\alpha}(-1)^{i-1}i\binom{\alpha}{i}\left(\sum_{s=0}^{\alpha-1} i^s c_s + (i-1)i^{\alpha-1}\right)u^{i-1}.$$

Hence by choosing the c's to satisfy the equations

$$\sum_{s=0}^{\alpha-1} i^s c_s + (i-1)i^{\alpha-1} = 0 \qquad (i = 1, 2, \cdots, \alpha) \tag{4.18}$$

as it is clear we may do, we see that the integrand in (4.13) vanishes identically, and the result for integral orders is established. But since the strength of the Cesàro method increases with the index, the general conclusion follows from this by Theorem 10.

One might be led to suspect that the preceding argument could be applied to obtain the desired conclusion when $q(u)$ is an arbitrary polynomial. It turns out that such is not the case, however, since in general the system of equations corresponding to (4.18) is inconsistent.

THEOREM 13. *The function $q(u) = u^p$ for $p > 0$ defines a perfect Hausdorff method.*

Proof. The normality and regularity are apparent. Let us then assume that (4.08) holds and replace therein $g_k(u)$ by its expression given in (4.07). Since $dq(u) = pu^{p-1}\,du$ we find on integrating by parts that

$$0 = (p+k)\int_0^1 u^k g_0^{(k)}(u)\,pu^{p-1}\,du = pg_0^{(k)}(1) - \int_0^1 u^{k+1}g_0^{(k+1)}(u)\,pu^{p-1}\,du.$$

Thus $g_0^{(k)}(1) = 0$ for $k = 0, 1, 2, \cdots$. From this the theorem follows.

It may be remarked that this result for $p \geqq 2$ is an immediate consequence of Corollary 2, and, for $p \geqq 1$, of the ensuing Theorem 15.

The next two theorems deal with certain general classes of monotone func-

tions. To facilitate the statement of the first, we introduce the following definition.

CONDITION D. Any function $q(u)$ defined as follows will be said to satisfy Condition D. For an arbitrarily given v, $0 < v \leqq 1$, let $U_1 \equiv (0 \leqq u < v)$ and $U_2 \equiv (v \leqq u \leqq 1)$. On U_1 let $q(u)$ be monotone increasing with $q(0) = 0$, and suppose that r exists, $0 \leqq r < 1$, such that $q(v - 0) \leqq r/(1 + r)$. Finally, at each point of U_2 let $q(u)$ be equal to 1.

THEOREM 14. *If $q(u)$ satisfies Condition* D, *then the corresponding Hausdorff matrix is of type M.*

Proof. For each $i = 1, 2, 3, \cdots$ let $0 \equiv u_0^i < u_1^i < \cdots < u_{m_i+1}^i \equiv v$ be a mode of subdividing the interval $0 \leqq u \leqq v$ such that the maximum length of the subdivisions tends to zero as $i \to \infty$. Let $Q_s^i \equiv q(u_s^i) - q(u_{s-1}^i)$ for $s = 1, 2, \cdots, m_i + 1$.

Assuming that (4.10) holds, we set

$$G_i(t) \equiv \sum_{s=1}^{m_i+1} g_0(tu_s^i)Q_s^i \qquad (0 \leqq t \leqq 1), \tag{4.19}$$

and we have then for each t, $\lim_{i\to\infty} G_i(t) = 0$. By Condition D,

$$Q_{m_i+1}^i \equiv 1 - q(u_{m_i}^i) \geqq 1/(1 + r),$$

$$0 \leqq \sum_{s=1}^{m_i} Q_s^i/Q_{m_i+1}^i \equiv q(u_{m_i}^i)/(1 - q(u_{m_i}^i)) \leqq r < 1,$$

for $i = 1, 2, 3, \cdots$. Consequently, if we let $g_0 \equiv \max | g_0(u) |$ on $(0, v)$, we obtain from (4.19) the inequality $| g_0(vt) | \leqq | G_i(t)/Q_{m_i+1}^i | + rg_0$. Letting $i \to \infty$ gives $| g_0(vt) | \leqq rg_0$ for $0 \leqq vt \leqq v$. This is not possible unless $g_0 = 0$. Thus $g_0(u) \equiv 0$, and the theorem is proved.

It is evident that no essential change is necessary in the above proof if $v < 1$ and $U_1 \equiv (0 \leqq u \leqq v)$, $U_2 \equiv (v < u \leqq 1)$.

THEOREM 15. *If $q(u)$ is continuous on U and has a derivative for $0 < u < 1$ which is non-negative and non-decreasing, then the corresponding Hausdorff matrix is of type M.*

Proof.[9] Under the given conditions $q(u)$ is absolutely continuous and (4.10) may be written

$$\int_0^1 g_0(tu)q'(u)\,du = 0 \qquad (0 \leqq t \leqq 1). \tag{4.20}$$

If we assume that $g_0(u)$ is not identically zero, it follows that the function

$$h(u) \equiv \int_0^u g_0(t)\,dt \qquad (0 \leqq u \leqq 1) \tag{4.21}$$

[9] The proof given here parallels an argument ascribed to E. J. McShane, which applies directly to the function $q(u) = (2/\pi) \sin^{-1} u$. See Bonnesen und Fenchel, *Theorie der Konvexen Körper*, p. 138. I am indebted to Professor Hans Lewy for calling this to my attention.

(which may be continued analytically into the circle $|u - 1| < 1$) must assume both positive and negative values. For integrating (4.20) with respect to t from 0 to v $(0 \leqq v \leqq 1)$ and interchanging the order of integration we obtain

$$\int_0^1 \frac{h(vu)}{u} q'(u)\, du = 0 \qquad (0 \leqq v \leqq 1). \tag{4.22}$$

Since $q(1) - q(0) = 1$, $q'(u)$ cannot be zero almost everywhere and hence the assumption that $h(u)$ is of one sign or zero implies that $h(u)$ is zero on some set of positive measure. This is a contradiction to (4.21) unless $g_0(u)$ is identically zero.

Let us denote then by $-m$ the minimum value of the function $h(t)/t$, which from (4.21) is continuous for $0 \leqq t \leqq 1$, if the value 0 is assigned at $t = 0$. Suppose that $h(t_0)/t_0 = -m$, where $0 < t_0 \leqq 1$, and set $p(t) \equiv h(t) + mt$ so that $p(t) \geqq 0$, $p(t_0) = 0$. Then for the function

$$F(t) \equiv \int_0^1 \frac{p(tu)}{u} q'(u)\, du = \int_0^1 \frac{h(tu)}{u} q'(u)\, du + mt \int_0^1 q'(u)\, du$$

we get from (4.22), $F(t) = mt(q(1) - q(0)) = mt$, and thus

$$F(t) - F(t_0) = m(t - t_0), \qquad m > 0. \tag{4.23}$$

On the other hand, if $t_0 < 1$ and $t_0 < t \leqq 1$ a simple calculation gives

$$F(t) - F(t_0) = \int_0^{t_0/t} \frac{p(tu)}{u}\left(q'(u) - q'\left(\frac{tu}{t_0}\right)\right) du + \int_{t_0/t}^1 \frac{p(tu)}{u} q'(u)\, du.$$

The first integral in view of $p(t) \geqq 0$ and the monotone property of $q'(u)$ is less than or equal to zero. Furthermore, for $t_0/t \leqq u \leqq 1$ it follows from (4.21) that

$$p(tu) = p(tu) - p(t_0) = h(tu) - h(t_0) + m(tu - t_0) \leqq (\max |g_0| + m)(t - t_0).$$

This shows that the second integral is $o(t - t_0)$. These estimates provide a contradiction to (4.23) and complete the proof in case $t_0 < 1$. A similar argument applies when $t_0 = 1$.

We conclude our discussion of the Hausdorff method by showing, as previously mentioned, that there exists a regular Hausdorff matrix which is not of type M.

Example 3. The polynomial $Q(u) = 16u^3 - 27u^2 + 12u$ defines a regular Hausdorff matrix which is not of type M. Since $Q(1) - Q(0) = 1$, the regularity is clear. Now let us choose $\alpha_0 = 0$, $\alpha_1 = -1$, and $\alpha_n = 1/(n(n-1))$ for $n \geqq 2$. Then we have $\sum_{n=0}^{\infty} |\alpha_n| < \infty$, $g_0(u) \equiv \sum_{n=0}^{\infty} \alpha_n(1-u)^n = u \log u$, and (4.07) gives $g_1(u) = -u(1 + \log u)$, $g_k(u) = u/(k(k-1))$ for $k \geqq 2$. Consequently, conditions (4.08) reduce simply to the two conditions $\int_0^1 u\, dQ(u) = 0$, $\int_0^1 (u \log u)\, dQ(u) = 0$, and one easily verifies that these are satisfied. Thus the matrix defined by $Q(u)$ is not of type M. Moreover,

since the diagonal elements of this matrix are given by

$$h_{nn} \equiv \int_0^1 u^n \, dQ(u) = 6(n-1)^2/((n+1)(n+2)(n+3)) \quad (n = 0, 1, 2, \cdots),$$

we see that the normality is destroyed by (and only by) the vanishing of h_{11}.

5. **The weighted-mean method.** The weighted-mean method is defined by a triangular matrix whose elements have the form $a_{nk} \equiv p_k/P_n$ for $k = 0, 1, 2, \cdots, n; n = 0, 1, 2, \cdots$, where the sequence $\{p_k\}$ is such that $P_n \equiv p_0 + p_1 + \cdots + p_n \neq 0$ for all n. If such a matrix is normal, we must have $a_{nn} \equiv p_n/P_n \neq 0$, or $p_n \neq 0$ for every n. In this event, conditions (1.4) reduce to $\sum_{n=k}^{\infty} \alpha_n/P_n = 0$ $(k = 0, 1, 2, \cdots)$, which clearly imply $\alpha_n = 0$ $(n = 0, 1, 2, \cdots)$. Thus we have established the following theorem.

THEOREM 16. *Every normal weighted-mean matrix is of type M.*

BROWN UNIVERSITY AND MICHIGAN STATE COLLEGE.

QUASI-UNITARY MATRICES

By John Williamson

Introduction. Let I_m be the n-rowed square matrix

$$\begin{pmatrix} E_m & 0 \\ 0 & -E_{n-m} \end{pmatrix},$$

where E_j is the unit matrix of order j. Then I_m is the normal form of a non-singular Hermitian matrix of index m under a non-singular conjunctive transformation. A matrix A, whose elements are complex numbers, which satisfies

$$AI_mA^* = I_m, \tag{1}$$

where $A^* = \bar{A}'$ is the conjugate transposed of A, will be called a *quasi-unitary matrix*. In particular, if $m = n$ or 0, A is a unitary matrix. A matrix, A, which satisfies (1), is a conjunctive automorph of the Hermitian matrix I_m. The conjunctive automorphs of a non-singular Hermitian matrix have been studied by Loewy.[1] He has shown how the nature of the elementary divisors of $A - \lambda E$ is restricted by the index m of the matrix I_m. In the following paper we derive normal forms for quasi-unitary matrices under quasi-unitary transformations, and in doing so are inevitably led to Loewy's results. (See, for example, the remark following Theorem 2.) We also determine necessary and sufficient conditions for the similarity of two quasi-unitary matrices under a quasi-unitary transformation. In particular it is shown that two quasi-unitary matrices which are similar are not necessarily similar under a quasi-unitary transformation. In §2 the similar problem for real quasi-orthogonal matrices is considered, and in §4 an interesting property of the elementary divisors of a pencil, whose base is I_m and a canonical quasi-unitary matrix, is deduced.

As many of the proofs are in essence the same, subject to obvious modifications, as those in a previous paper,[2] for the sake of brevity they will be omitted.

Received May 4, 1937; in revised form September 21, 1937.

[1] Alfred Loewy, *Allgemeine bilineare Formen konjugirt imaginären Variabeln*, Abhandlungen der Kaiserlichen Leopoldinisch-Carolinischen Deutschen Akademie der Naturforscher, vol. 71 (1898), pp. 377–446; Mathematische Annalen, vol. 50, pp. 557–576. The second of these papers gives a short account of the results proved in the first. The term quasi-unitary was first used by Harold Hilton, *Properties of certain homogeneous linear substitutions*, Annals of Mathematics, (2), vol. 15 (1913), pp. 195–201.

[2] John Williamson, *On the normal forms of linear canonical transformations in dynamics*, American Journal of Mathematics, vol. 59 (1937), pp. 599–617. This paper will be referred to as W.

1. The problem that we first consider, then, is the following. Let A_1 and A_2 be two matrices, which satisfy

$$A_i I_m A_i^* = I_m \qquad (i = 1, 2, 3); \tag{2}$$

to determine necessary and sufficient conditions that a third matrix A_3, satisfying (2), exist and satisfy $A_3 A_2 A_3^{-1} = A_1$.

If A_1 and A_2 are similar and a matrix Q, to be specified later, is similar to A_1, then Q is also similar to A_2. There accordingly exist two non-singular matrices R_1 and R_2, such that

$$R_i A_i R_i^{-1} = Q \qquad (i = 1, 2).$$

The matrices

$$S_i = R_i I_m R_i^* \qquad (i = 1, 2)$$

are Hermitian and are left invariant by Q; that is, they satisfy

$$Q S_i Q^* = S_i \qquad (i = 1, 2).$$

Accordingly, if Q is any matrix similar to both of the quasi-unitary matrices A_1 and A_2, there is *associated with* A_1 a Hermitian matrix S_1 and with A_2 a Hermitian matrix S_2, both of which are left invariant by Q.

The problem under consideration is reduced to a similar but simpler one by means of

THEOREM 1. *A necessary and sufficient condition that the quasi-unitary matrix A_1 be similar to the quasi-unitary matrix A_2 under a quasi-unitary transformation is that there exist a non-singular matrix H, such that*

$$HQ = QH,$$

and that the two Hermitian matrices associated with A_1 and A_2 satisfy[3]

$$HS_1H^* = S_2.$$

Since, in the above, Q is any matrix similar to A_1, we are at liberty to choose Q in a suitable normal form. Then, if S is any Hermitian matrix, which satisfies the equation

$$QSQ^* = S, \tag{3}$$

we shall first determine a unique normal form for S under non-singular conjunctive transformations by matrices commutative with Q. If $HQ = QH$ and $HSH^* = T$, we shall call the transformation by the matrix H an *admissible transformation* and shall write $S = T$.

Let the matrix Q, which is a normal form of the quasi-unitary matrix A under similarity transformations, be chosen in the diagonal block form

$$Q = [Q_1, Q_2, \cdots, Q_k],$$

[3] For proof, see W, Theorem 1.

where no latent root of Q_1 has absolute value one, each latent root of Q_j, $j > 1$, has absolute value one, and, if $i \neq j$, no latent root of Q_i is the same as a latent root of Q_j. Then, if (3) is satisfied, the matrix S is also a diagonal block matrix,

$$S = [S_1, S_2, \cdots, S_k]$$

and

$$Q_i S_i Q_i^* = S_i \qquad (i = 1, 2, \cdots, k) \tag{4}$$

(W, Lemma 2). Since (4) is the same as (3) except for the suffix i, we need only consider two special cases of Q:

Case 1. No latent root of Q has absolute value one.

Case 2. Each latent root of Q is equal to p, where p is of absolute value one.

Case 1. Since Q is similar to $(Q^*)^{-1}$, Q is similar to the diagonal block matrix $[F, (F^*)^{-1}]$. As a consequence of the remark following Theorem 1, we may replace Q by this matrix; that is, we may write $Q = [F, (F^*)^{-1}]$. With this value of Q, the matrix S is of the form

$$\begin{pmatrix} 0 & T \\ T^* & 0 \end{pmatrix},$$

where T is a square matrix of the same order as F. The transformation of matrix

$$\begin{pmatrix} T^{-1} & 0 \\ 0 & E \end{pmatrix}$$

is admissible and

$$S = \begin{pmatrix} 0 & E \\ E & 0 \end{pmatrix} = G. \tag{5}$$

Hence we have

RESULT 1. *If no latent root of A has absolute value one, the matrix Q may be taken in the form $[F, (F^*)^{-1}]$. Then $S = G$.*

The matrix F is not unique and may be replaced by any matrix similar to it, the classical canonical form, for instance. As a consequence of Theorem 1 we therefore have

THEOREM 2. *If A_1 is a quasi-unitary matrix similar to a second quasi-unitary matrix A_2 and, if no latent root of A_1 is of absolute value one, then A_1 is similar to A_2 under a quasi-unitary transformation.*

The fact that in this case the index of I_m must be one half the order of I_m is a known result.[4]

[4] Alfred Loewy, loc. cit.

Case 2. Since each latent root of Q is equal to p, we may take Q in the canonical form

$$Q = [P_{e_1}, P_{e_2}, \cdots, P_{e_t}],$$

where

(6) $$P_{e_j} = pE_j + pU_j,$$

and E_j and U_j are respectively the unit matrix and the auxiliary unit matrix of order e_j. The elementary divisors of $A - \lambda E$ are therefore

$$(\lambda - p)^{e_i} \qquad (i = 1, 2, \cdots, t; e_1 \geqq e_2 \geqq \cdots \geqq e_t).$$

If

(7) $$S = (S_{rs}) \qquad (r, s = 1, 2, \cdots, t),$$

is a partition of S similar to that of Q, it can be shown first that $S \approx T = (T_{rs})$, where T_{11} is non-singular (W, §4); then that $T \approx [S_1, S_2, \cdots, S_t]$, where

(8) $$P_{e_j} S_j P^*_{e_j} = S_j$$

(W, Lemma 3). Equations (8) are of two distinct types: *type* (1), the matrix $P_{e_j} = P$ is of even order $2m$; *type* (2), the matrix P is of odd order $e = 2m + 1$.

Type (1). The reduction, used in W, type b, shows that

$$S_j \approx W_j = d_j X_j = d_j \begin{pmatrix} 0 & X_{12} \\ X_{21} & 0 \end{pmatrix},$$

where X_{12} is a uniquely determined square matrix, all of whose elements are integers, and $X_{21} = -X'_{12}$. For example,[5] if $m = 4$,

$$X_{12} = \begin{Vmatrix} 1 & 3 & 3 & 1 \\ -1 & -2 & -1 & 0 \\ 1 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{Vmatrix}.$$

Since $W_j = W^*_j$, $d_j = \alpha_j i$, where α_j is a real number different from zero. The admissible transformation by the scalar matrix $E/\sqrt{|\alpha_j|}$ shows that $S_j \approx W_j = \epsilon_j i X_j$, where $\epsilon_j = \pm 1$. Therefore we have

Result 2. *If* $e_j = 2m$, $S_j \approx \epsilon_j i X_j$, *where* $\epsilon_j = \pm 1$ *and* X_j *is uniquely determined.*

Type (2). The matrix $S_j \approx \epsilon_j Y_j$, $\epsilon_j = \pm 1$, where $Y_j = (y_{rs})$ is a uniquely determined matrix, for which

$$y_{rs} = 0 \qquad (r, s = 1, 2, \cdots, m; y_{rs} = 0, r + s \geqq e_j + 2),$$

[5] Cf. Turnbull and Aitken, *Canonical Matrices*, p. 157.

(W, type b_2). For example, if $e_i = 5$,

$$Y_i = \begin{pmatrix} 0 & 0 & \frac{1}{2} & \frac{3}{2} & 1 \\ 0 & 0 & -\frac{1}{2} & -1 & 0 \\ \frac{1}{2} & -\frac{1}{2} & 1 & 0 & 0 \\ \frac{3}{2} & -1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Consequently, we have

RESULT 3. *If* $e_i = 2m + 1$, $S_i \approx \epsilon_i Y_i$, *where* $\epsilon_i = \pm 1$ *and* Y_i *is uniquely determined.*

We see from Results 2 and 3 that, with each elementary divisor $(\lambda - p)^e$ of $A - \lambda E$, where $|p| = 1$, is associated an ϵ, which has the value ± 1. Therefore, if $(\lambda - p)^e$ occurs exactly t times among the elementary divisors of $A - \lambda E$, with this elementary divisor is associated a set of t positive and negative signs. We may call the number of these positive signs the *index of the elementary divisor* $(\lambda - p)^e$. We are now able to prove (W, Theorem 4)

THEOREM 3. *Necessary and sufficient conditions that two quasi-unitary matrices* A_1 *and* A_2 *be equivalent under a quasi-unitary transformation are that*

(a) *the elementary divisors of* $A_1 - \lambda E$ *be the same as those of* $A_2 - \lambda E$, *and*

(b) *the indices of all elementary divisors* $(\lambda - p)^e$, $|p| = 1$, *be the same for both pencils.*

Theorem 3 includes as a special case the known theorem that two unitary matrices which are similar are similar under a unitary transformation. For, if A is unitary, only elementary divisors of type (2) may occur with $e_i = 1$ and the corresponding indices must all be one (or zero).

2. **Real quasi-orthogonal matrices.** The above arguments are valid in the real field, if the complex number i is replaced by the two-rowed real matrix

$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

and the complex number $p = a + ib$ of unit modulus by the real orthogonal matrix

$$\begin{pmatrix} a & b \\ -b & a \end{pmatrix}.$$

The elementary divisors $(\lambda \pm 1)^{e_i}$ of type (1), $e_i = 2m$, now must be considered separately. In this case, since the matrix S_{11} in (7) is a symmetric matrix of even order, all of whose elements are real numbers, S_{11} is necessarily singular.

However, after at most a rearrangement of the rows and columns of S, we may suppose that

$$\begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix}$$

is non-singular. Then $[P, P]$ may be replaced by $[P, (P^*)^{-1}]$ and

$$\begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix} \approx \begin{pmatrix} 0 & E \\ E & 0 \end{pmatrix}$$

(W, type a). Accordingly, an elementary divisor $(\lambda \pm 1)^{2m}$ must occur an even number of times and no index need be associated with it. We have therefore

THEOREM 4. *Two real quasi-orthogonal matrices A_1 and A_2 are similar under a real quasi-orthogonal transformation, if and only if*

(a) *the elementary divisors of $A_1 - \lambda E$ are the same as those of $A_2 - \lambda E$, and*

(b) *the indices associated with each pair of complex elementary divisors $(\lambda - p)^k$, $(\lambda - \bar{p})^k$, $|p| = 1$, and with each elementary divisor $(\lambda \pm 1)^{2k+1}$ are the same for both pencils.*

3. **Normal Forms.** In determining possible normal forms for a quasi-unitary matrix under quasi-unitary transformations we first reduce the matrices X_j and Y_j of Results 2 and 3 to simpler forms. In so doing we naturally alter the matrices P_{e_j}.

Type (1). Since $e_j = 2m$, we may write

$$P_{2m} = \begin{pmatrix} P_m & R_m \\ 0 & P_m \end{pmatrix},$$

where R_m is a square matrix of order m, whose only non-zero element is the element p in the first column and last row. Then, if

$$H = \begin{pmatrix} E_m & 0 \\ 0 & -i\epsilon X_{21}^{-1} \end{pmatrix}, \text{ where } \epsilon_j = \epsilon,$$

$$H\epsilon iXH^* = \begin{pmatrix} 0 & E_m \\ E_m & 0 \end{pmatrix} = G_m \quad \text{and} \quad HP_{2m}H^{-1} = \begin{pmatrix} P_m & -\epsilon iL_m \\ 0 & (P_m^*)^{-1} \end{pmatrix},$$

where L_m is the matrix, whose last row is

$$(p, -p, p, \cdots, (-1)^{m-1}p), \tag{9}$$

all other rows being zero. Therefore, since $\epsilon = \pm 1$, we have

RESULT 2a. *If $e_j = 2m_j$, P_{2m_j} may be replaced by*

$$Z_{2m_j} = \begin{pmatrix} P_{m_j} & \epsilon iL_{m_j} \\ 0 & (P_{m_j}^*)^{-1} \end{pmatrix}.$$

Then $S_{2m_j} \approx G_{m_j}$.

Type (2). Since $e_j = 2m + 1$, we may write

$$P_{2m+1} = \begin{pmatrix} P_m & R_m \\ 0 & P_{m+1} \end{pmatrix},$$

where R_m is a matrix of m rows and $m + 1$ columns, whose only non-zero element is the element p in the last row and first column. The corresponding matrix Y in Result 2 is of the form

$$Y = \begin{pmatrix} 0 & C \\ D & 0 \end{pmatrix},$$

where D is a non-singular $(m + 1)$-rowed matrix and C consists of the first m rows of D^*. If

$$H = \begin{pmatrix} E_m & 0 \\ 0 & \epsilon D^{-1} \end{pmatrix},$$

$$H\epsilon YH^* = [G_m, \epsilon(-1)^m] \qquad \text{and} \qquad HP_{2m+1}H^{-1} = \begin{pmatrix} P_m & M_m \\ 0 & (P^*_{m+1})^{-1} \end{pmatrix},$$

where $M_m = \epsilon R_m D$. The first $m - 1$ rows of M_m are therefore zero and the last is

$$(\tfrac{1}{2}\epsilon p, -\tfrac{1}{2}\epsilon p, \tfrac{1}{2}\epsilon p, \cdots, (-1)^{m-1}\tfrac{1}{2}\epsilon p, (-1)^m\epsilon p). \tag{10}$$

On substituting $(-1)^m\epsilon$ for ϵ we have

RESULT 3a. *If $e_j = 2m_j + 1$, P_{e_j} may be replaced by*

$$Z_{e_j} = \begin{pmatrix} P_{m_j} & (-1)^{m_j} M_{m_j} \\ 0 & (P^*_{m_j+1})^{-1} \end{pmatrix}.$$

Then $S_{e_j} = [G_{m_j}, \epsilon]$.

If O is the real orthogonal matrix

$$O = 2^{-\frac{1}{2}}\begin{pmatrix} E & E \\ -E & E \end{pmatrix},$$

then

$$OGO' = [E, -E]. \tag{11}$$

Further, if $Z = [F, (F^*)^{-1}]$, where F is the matrix of Result 1,

$$OZO' = B = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix} = \tfrac{1}{2}\begin{pmatrix} F + (F^*)^{-1} & -F + (F^*)^{-1} \\ -F + (F^*)^{-1} & F + (F^*)^{-1} \end{pmatrix}, \tag{12}$$

and, if Z has the value given in Result 2a,

$$OZO' = B = \tfrac{1}{2}\begin{pmatrix} P + (P^*)^{-1} + \epsilon iL & -P + (P^*)^{-1} + \epsilon iL \\ -P + (P^*)^{-1} - \epsilon iL & P + (P^*)^{-1} - \epsilon iL \end{pmatrix}. \tag{13}$$

In reducing the matrix $[G_{m_j}, \epsilon]$ of Result 3a to diagonal form it is necessary to make a further partition of the matrix Z_{e_j}. Accordingly we write

$$P_{m_j} = P, \qquad (-1)^{m_j} M_j = Nx, \qquad (P^*_{m_j+1})^{-1} = \begin{pmatrix} (P^*)^{-1} & 0 \\ y^* & (p^*)^{-1} \end{pmatrix},$$

where N is an m-rowed square matrix, x a matrix of a single column, and y^* a matrix of a single row. An easy calculation shows that

$$(14) \qquad [O, 1][G, \epsilon][O, 1]' = [E_m, -E_m, \epsilon],$$

and that

$$(15) \qquad [O, 1]Z[O, 1]' = B = (B_{ij}) \qquad (i, j = 1, 2),$$

where

$$B_{11} = \{P + N + (P^*)^{-1}\}2^{-1}, \qquad B_{12} = (\{-P + (P^*)^{-1} + N\}2^{-1}, \quad x2^{-\frac{1}{2}}),$$

$$B_{21} = \begin{pmatrix} \{-P - N + (P^*)^{-1}\}2^{-1} \\ y^*2^{-\frac{1}{2}} \end{pmatrix}, \qquad B_{22} = \begin{pmatrix} \{P + (P^*)^{-1} - N\}2^{-1}, & -x2^{-\frac{1}{2}} \\ y^*2^{-\frac{1}{2}}, & (p^*)^{-1} \end{pmatrix}.$$

If $\epsilon = -1$, the matrix on the right of (14) is $[E_m, -E_{m+1}]$. If $\epsilon = +1$, a simple interchange of rows and the same interchange of columns reduces the matrix on the right of (14) to $[E_{m+1}, -E_m]$. Accordingly, if $\epsilon = 1$, there exists a real orthogonal matrix O_1 such that

$$O_1[G, 1]O_1' = [E_{m+1}, -E_m],$$

and

$$(16) \qquad O_1 Z O_1' = B = (B_{ij}) \qquad (i, j = 1, 2),$$

where

$$B_{11} = \begin{pmatrix} \{P + N + (P^*)^{-1}\}2^{-1}, & x2^{-\frac{1}{2}} \\ y^*2^{-\frac{1}{2}}, & (p^*)^{-1} \end{pmatrix}, \qquad B_{12} = \begin{pmatrix} \{-P + N + (P^*)^{-1}\}2^{-1} \\ y^*2^{-\frac{1}{2}} \end{pmatrix},$$

$$B_{21} = (\{-P - N + (P^*)^{-1}\}2^{-1}, \quad -x2^{-\frac{1}{2}}), \qquad B_{22} = \{P - N + (P^*)^{-1}\}2^{-1}.$$

Thus each matrix Z_j in Results 1, 2a and 3a is similar under a real orthogonal transformation to a matrix B_j given by one of the equations (12), (13), (15) or (16) and the corresponding matrix $S_j = [E_t, -E_s]$. Let $B_1, B_2, \cdots, B_k$ be the complete set of matrices B_i, described above, obtained from a quasi-unitary matrix A, and let

$$B_r = (B_{r,ij}) \qquad (i, j = 1, 2;\ r = 1, 2, \cdots, k).$$

Then, if

$$(17) \qquad C = (C_{ij}) \qquad (i, j = 1, 2),$$

where

$$(18) \qquad C_{ij} = [B_{1,ij}, B_{2,ij}, \cdots, B_{k,ij}],$$

it follows that A is similar to C and that $S = I_m$. We have therefore proved

THEOREM 5. *A quasi-unitary matrix A is similar under a quasi-unitary transformation to one and essentially only one of the matrices C defined by* (17) *and* (18).

4. Elementary divisors. We now prove three lemmas.

LEMMA 1. *The elementary divisors of $[P_m, (P_m^*)^{-1}] - \lambda G_m$ are all linear and of the form $\lambda - \omega$, where ω is of absolute value one.*

Let

$$\Delta = \begin{vmatrix} P_m & -\lambda E_m \\ -\lambda E_m & (P_m^*)^{-1} \end{vmatrix}.$$

Then[6]

$$\Delta = | P_m(P_m^*)^{-1} - \lambda^2 E_m |.$$

But, if $\alpha = p/\bar{p} = e^{2i\theta}$, it follows from (6) that

$$P_m(P_m^*)^{-1} = \begin{bmatrix} 0 & \alpha & 0 & \cdots & 0 \\ 0 & 0 & \alpha & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & \alpha \\ (-1)^{m-1}\alpha & (-1)^{m-2}\alpha & (-1)^{m-3}\alpha & \cdots & \alpha \end{bmatrix}.$$

Therefore, if $\lambda = \mu^{\frac{1}{2}} e^{i\theta}$,

$$\begin{aligned} \Delta &= (-1)^m \alpha^m [\mu^m - \mu^{m-1} + \mu^{m-2} - \cdots + (-1)^m], \\ &= (-1)^m \alpha^m (\mu^{m+1} + (-1)^m)/(\mu + 1). \end{aligned}$$

The roots of $\Delta = 0$ are accordingly all distinct and of absolute value one and the lemma is proved.

We have as an immediate

COROLLARY. *The matrix F of Result* 1 *may be so chosen, that the elementary divisors of $[F, (F^*)^{-1}] - \lambda G$ are all linear and of the form $\lambda - \omega$, where ω is of absolute value one.*

LEMMA 2. *If Z_m is the matrix of Result 2a, the elementary divisors of $Z_m - \lambda G_m$ are all linear and of the form $\lambda - \omega$, where ω is of absolute value one.*

If $\Delta = | Z_m - \lambda G_m |$, then[6] $\Delta = | P_m(P_m^*)^{-1} + \lambda \epsilon i L_m - \lambda^2 E_m |$.

Since p is of absolute value 1, $p = e^{i\theta}$ and $(p^*)^{-1} = p$. Hence on substituting for L_m its value given by (9) and on writing $\lambda = (-1)^{m-1} \epsilon i e^{i\theta} \mu$ and $f = 1 + \mu$, we find that

$$\begin{aligned} \Delta &= e^{2mi\theta} \begin{vmatrix} \mu^2 & 1 & 0 & \cdots & 0 \\ 0 & \mu^2 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ (-1)^{m-1}f & (-1)^{m-2}f & (-1)^{m-3}f & \cdots & \mu^2 + f \end{vmatrix} \\ &= e^{2mi\theta}(\mu^{2m} + \mu^{2m-1} + \mu^{2m-2} + \cdots + \mu + 1). \end{aligned}$$

Consequently the roots of $\Delta = 0$ are all distinct and of absolute value one.

[6] J. Williamson, *The expansion of determinants of composite order*, American Mathematical Monthly, vol. 40 (1933), p. 67, formula 7.

LEMMA 3. *If Z is the matrix of Result 3a, the elementary divisors of $Z - \lambda[G_m, \epsilon]$ are all linear and of the form $\lambda - \omega$, where ω is of absolute value one.*

If $p = e^{i\theta}$ and Δ is the determinant of this pencil, we deduce, as in Lemmas 1 and 2, that Δ is equal to a determinant of order $m + 1$; in fact

$$\Delta = \begin{vmatrix} -\lambda^2 & e^{2i\theta} & \cdots & 0 & 0 \\ 0 & -\lambda^2 & \cdots & 0 & 0 \\ (-1)^{m-1}\beta & (-1)^{m-2}\beta & \cdots & \beta - \lambda^2 & \epsilon\lambda e^{i\theta} \\ (-1)^m e^{i\theta} & (-1)^{m-1}e^{i\theta} & \cdots & -e^{i\theta} & e^{i\theta} - \lambda\epsilon \end{vmatrix},$$

where $\beta = e^{2i\theta} - \epsilon e^{i\theta}\lambda/2$. On writing $\epsilon\lambda = \mu e^{i\theta}$ and adding β times the last row to the last but one, we see that

$$(19) \qquad \Delta = e^{(2m+1)i\theta} \begin{vmatrix} -\mu^2 & 1 & \cdots & 0 & 0 \\ 0 & -\mu^2 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & -\mu^2 & 1 - \frac{1}{2}\mu + \frac{1}{2}\mu^2 \\ (-1)^m & (-1)^{m-1} & \cdots & -1 & 1 - \mu \end{vmatrix}.$$

On replacing the last row of the determinant on the right of (18) by $(-1)^m\text{row}_1 + (-1)^{m-1}\text{row}_2 + \cdots - \text{row}_m + (\mu^2 + 1)\text{row}_{m+1}$, we obtain

$$(1 + \mu^2)\Delta = e^{(2m+1)i\theta} \begin{vmatrix} -\mu^2 & 1 & \cdots & 0 & 0 \\ 0 & -\mu^2 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & -\mu^2 & 1 - \frac{1}{2}\mu + \frac{1}{2}\mu^2 \\ (-1)^m & 0 & \cdots & 0 & \gamma \end{vmatrix} = e^{(2m+1)i\theta}\phi(\mu),$$

where $\gamma = (1 + \mu^2)(1 - \mu) - (1 - \frac{1}{2}\mu + \frac{1}{2}\mu^2) = \mu(\frac{1}{2}\mu - \frac{1}{2} - \mu^2)$. Hence

$$\begin{aligned} \phi(\mu) &= (-1)^m\gamma\mu^{2m} + 1 - \tfrac{1}{2}\mu + \tfrac{1}{2}\mu^2, \\ &= (-1)^{m+1}\mu^{2m+1}[\mu^2 - \tfrac{1}{2}\mu + \tfrac{1}{2}] + \tfrac{1}{2}\mu^2 - \tfrac{1}{2}\mu + 1. \end{aligned}$$

We now proceed to show that the equation

$$(20) \qquad \phi(\mu) = 0,$$

has $2m + 3$ distinct roots of absolute value one. If m is odd and $\mu = e^{2it}$, (19) is equivalent to

$$(21) \qquad f(t) = \cos(2m + 3)t - \tfrac{1}{2}\cos(2m + 1)t + \tfrac{1}{2}\cos(2m - 1)t = 0.$$

Let

$$t_k = \frac{k\pi}{2m + 3} \qquad (k = 0, 1, 2, \cdots, 2m + 3).$$

If k is odd, $\cos(2m+3)t_k = -1$, while $|\cos(2m+1)t_k| < 1$ and

$$|\cos(2m-1)t_k| < 1.$$

Accordingly, when k is odd, $f(t_k)$ is negative. Similarly, when k is even, $f(t_k)$ is positive. Consequently, $f(t)$ has one zero between t_k and t_{k+1} and therefore $f(t)$ has at least $2m+3$ distinct zeros between 0 and π. Hence there are at least $2m+3$ distinct values of t between 0 and π, such that e^{2it} is a zero of $\phi(\mu)$. There are therefore $2m+3$ distinct roots of $\phi(\mu)=0$ which have absolute value one, and, since $\phi(\mu)$ is of degree $2m+3$, all the roots of $\phi(\mu)=0$ are distinct and of absolute value one. By a slight modification of the above argument we arrive at the same result when m is even. Therefore the roots of $\phi(\mu)=0$ are all distinct and have absolute value one. Since the final reduction to normal form in §4 was orthogonal, it is a consequence of Lemmas 1, 2 and 3 that, if C is the matrix defined by (17), the elementary divisors of $C-\lambda I_m$ are all linear and of the form $\lambda-\omega$, where ω is of absolute value one. Accordingly we have proved

THEOREM 6. *If A is a quasi-unitary matrix (unitary with respect to I_m), then A is similar under a quasi-unitary transformation to a matrix C, where the elementary divisors of $C-\lambda I_m$ are all simple and of the form $\lambda-\omega$, where ω is of absolute value one.*

If A is unitary, that is, if $I_m = E$, Theorem 6 remains true when C is replaced by A, and we obtain the theorem that the latent roots of a unitary matrix are all of absolute value one. That no such simplification is possible for quasi-unitary matrices is shown by the following example.

If F is an arbitrary non-singular matrix, then

$$\begin{pmatrix} 0 & F \\ (F^*)^{-1} & 0 \end{pmatrix}\begin{pmatrix} 0 & E \\ E & 0 \end{pmatrix}\begin{pmatrix} 0 & F^{-1} \\ F^* & 0 \end{pmatrix} = \begin{pmatrix} 0 & E \\ E & 0 \end{pmatrix},$$

and the elementary divisors of

$$\begin{pmatrix} 0 & F \\ (F^*)^{-1} & 0 \end{pmatrix} - \lambda\begin{pmatrix} 0 & E \\ E & 0 \end{pmatrix}$$

are the elementary divisors of $F-\lambda E$ together with those of $(F^*)^{-1}-\lambda E$. If O is the orthogonal matrix which reduces

$$\begin{pmatrix} 0 & E \\ E & 0 \end{pmatrix}$$

to I_m, the matrix

$$O\begin{pmatrix} 0 & F \\ (F^*)^{-1} & 0 \end{pmatrix}O' = A$$

is quasi-unitary, and the elementary divisors of $A-\lambda I_m$ are still the elementary divisors of $F-\lambda E$ and $(F^*)^{-1}-\lambda E$. Hence the elementary divisors of $A-\lambda I_m$ need not be simple or even of the type $(\lambda-\omega)^e$, where ω is of absolute value one.

THE JOHNS HOPKINS UNIVERSITY.

STABLE LAWS OF PROBABILITY AND COMPLETELY MONOTONE FUNCTIONS

By S. Bochner

In discussing stability of laws of probability other than the Gaussian, P. Lévy[1] has proved the following two statements about the Fourier transform

$$V_\rho(\alpha) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-ix\alpha} e^{-|x|^{2\rho}}\,dx.$$

I. If $0 < \rho \leqq 1$, $V_\rho(\alpha)$ is non-negative for all real values of α.

II. If $1 < \rho < \infty$, $V_\rho(\alpha)$ assumes in $-\infty < \rho < \infty$ both positive and negative values.

It is not hard to prove statement II. As for statement I, a simple proof is available in case $0 < \rho < \frac{1}{2}$, but this proof cannot be extended to cover the case $\frac{1}{2} < \rho < 1$.[2]

In the present note we shall give two new proofs for statement I. They are not of the easiest type, perhaps, but they do not distinguish between the two cases and they lead to more general classes of functions having non-negative Fourier transforms.

We shall consider the class $\mathfrak{P}$ of positive-definite functions

$$f(x) = \int_{-\infty}^{\infty} e^{i\alpha x}\,dV(\alpha),$$

which are Fourier transforms of bounded non-negative distributions. They have the following properties:

1. if $f_1, f_2 \subset \mathfrak{P}$, $a_1 \geqq 0$, $a_2 \geqq 0$, then $a_1 f_1 + a_2 f_2 \subset \mathfrak{P}$;
2. if $f_1, f_2 \subset \mathfrak{P}$, then $f_1 f_2 \subset \mathfrak{P}$;
3. if $f_n \subset \mathfrak{P}$ and $\lim_{n\to\infty} f_n$ exists uniformly in every finite interval, then $\lim f_n \subset \mathfrak{P}$.[3]

First proof. Excluding the trivial case $\rho = 1$, we have to prove that

$$f_\rho(x) = \exp\{-|x|^{2\rho}\}$$

belongs to $\mathfrak{P}$ for $0 < \rho < 1$. Since, for these values of ρ,

$$|x|^{2\rho} = c_\rho \int_0^\infty \frac{\alpha^{2\rho-1}\,d\alpha}{1 + \left(\dfrac{\alpha}{x}\right)^2}, \qquad c_\rho > 0,$$

Received September 20, 1937.

[1] P. Lévy, *Calcul des Probabilités*, 1925, pp. 252–277.

[2] G. Pólya, *Herleitung des Gaussschen Fehlergesetzes aus einer Funktionalgleichung*, Math. Zeitschrift, vol. 18 (1923), p. 109.

[3] P. Lévy, loc. cit., Chapitre II; S. Bochner, *Vorlesungen über Fouriersche Integrale*, 1932, pp. 63–77

$f_\rho(x)$ is, on every finite x-interval, the uniform limit of functions of the type

$$\exp\left[-\sum_{\nu=1}^{n} \frac{a_\nu^2}{1+\left(\frac{b_\nu}{x}\right)^2}\right].$$

Hence, by properties 3 and 1, it is sufficient to prove that

$$\exp\left\{-\frac{a^2x^2}{x^2+b^2}\right\} = e^{-a^2}\cdot\exp\left\{\frac{a^2b^2}{x^2+b^2}\right\}$$

belongs to $\mathfrak{P}$. But

$$\exp\left\{\frac{c^2}{x^2+b^2}\right\} = \sum_{n=0}^{\infty} \frac{c^{2n}}{n!}(x^2+b^2)^{-n}.$$

Therefore, because of properties 1, 2, 3, our assertion follows from the elementary fact that $(x^2+b^2)^{-1}$ belongs to $\mathfrak{P}$, namely,

$$(x^2+b^2)^{-1} = \frac{1}{2b}\int_{-\infty}^{\infty} e^{i\alpha x}e^{-b|\alpha|}\,d\alpha.$$

A general class of functions including $f_\rho(x)$, for which our argument remains in force, are

$$f_\varphi(x) = \exp\{-x^2\varphi(|x|)\},$$

where

$$\varphi(x) = \int_0^{\infty} \frac{d\gamma(\alpha)}{x^2+\alpha^2}, \qquad d\gamma(\alpha) \geqq 0.$$

These functions $\varphi(x)$ are a type of completely monotone functions which have been recently investigated by D. V. Widder and R. P. Boas.[4]

Second proof. We shall use the following lemma.[5] If a function $f(y)$ is completely monotone in $0 \leqq y < \infty$, and $\psi(y)$ is a function vanishing at the origin whose derivative $\psi'(y)$ is completely monotone in $0 < y < \infty$, then $f(\psi(y))$ is again completely monotone in $0 \leqq y < \infty$. Therefore,

$$f(\psi(y)) = \int_0^{\infty} e^{-yt}\,d\gamma(t), \qquad d\gamma(t) \geqq 0.$$

An admissible substitution is

$$\psi(y) = y^\rho, \qquad 0 < \rho < 1,$$

since

$$\psi'(y) = \rho y^{\rho-1} = \frac{\rho}{\Gamma(1-\rho)}\int_0^{\infty} e^{-yt}t^{-\rho}\,dt.$$

[4] D. V. Widder, *The iterated Stieltjes transform*, Proc. Nat. Acad. Sc., vol. 23 (1937), pp. 242–244.

[5] *Completely monotone functions of the Laplace operator for torus and sphere*, this Journal, vol. 3 (1937), pp. 488–503.

Putting $f(y) = e^{-y}$, we obtain

$$e^{-|y|^{\rho}} = \int_0^\infty e^{-yt}\, d\gamma(t),$$

or

$$e^{-|y|^{2\rho}} = \int_0^\infty e^{-y^2 t}\, d\gamma(t).$$

Obviously the right side is the uniform limit of finite sums of the type

$$\sum_{\nu=1}^{n} b_\nu^2 e^{-y^2 a_\nu^2}.$$

But $e^{-a^2 y^2}$ belongs to $\mathfrak{P}$ and therefore, by properties 1, 2, 3, the function $f_\rho(x)$ does also.

Princeton University.

ON SOME GENERALIZATIONS OF A THEOREM OF A. MARKOFF

BY EINAR HILLE, G. SZEGÖ, AND J. D. TAMARKIN

I. Introduction

1. The theorems of A. Markoff and of S. Bernstein concerning the derivative of a rational or of a trigonometric polynomial state that if $||f||$ denotes the maximum of the absolute value of a rational polynomial $f(x)$ over a finite interval (a, b), or of a trigonometric polynomial over its interval of periodicity, then for the derivative $f'(x)$ we have

$$(1.1)\qquad ||f'|| \leqq An^2 ||f|| \quad \text{or} \quad ||f'|| \leqq An ||f||,$$

respectively, where n is the degree of $f(x)$ and A is a constant which does not depend on n or on f, but only on $(b - a)$. In fact $A = 1$ in the case of rational polynomials considered on $(-1, +1)$ and also in the case of trigonometric polynomials of period 2π. These results can be stated in "abstract" form if we consider $f(x)$ as an element of the space C of continuous functions and interpret $||f||$ as the "norm" of this element. A natural question arises then whether estimates similar to (1.1) hold if $f(x)$ is considered as an element of other function spaces with different definition of the norm. The purpose of the present note is to answer this question for rational polynomials in the case of the space L_p, $p \geqq 1$, where the norm is defined by

$$||f|| = ||f||_p = \left\{\frac{1}{b - a}\int_a^b |f(x)|^p dx\right\}^{1/p}.$$

2. The corresponding problem for trigonometric polynomials was solved in a much more general case by Zygmund[1] by using an important interpolation formula of M. Riesz.[2] According to this formula we have, for an arbitrary trigonometric polynomial of degree n and of period 2π,

$$(1.2)\qquad |f'(x)| \leqq \sum_{\nu=1}^{2n} \rho_\nu^{(n)} |f(x + \theta_\nu^{(n)})|,$$

where $\rho_\nu^{(n)}$, $\theta_\nu^{(n)}$ are certain numbers which do not depend on $f(x)$ and which satisfy

$$(1.3)\qquad \rho_\nu^{(n)} > 0, \qquad \sum_{\nu=1}^{2n} \rho_\nu^{(n)} = n,$$

$$(1.4)\qquad 0 < \theta_1^{(n)} < \theta_2^{(n)} < \cdots < \theta_{2n}^{(n)} < 2\pi.$$

Received October 2, 1937.

[1] *A remark on conjugate functions*, Proceedings of the London Math. Soc., (2), vol. 34 (1932), pp. 392–400, esp. pp. 394–396.

[2] *Eine trigonometrische Interpolationsformel und einige Ungleichungen für Polynome*, Jahresbericht der Deutschen Math. Verein., vol. 22 (1914), pp. 354–368, esp. p. 356, (9), (10).

Now let $\phi(u)$ be any non-negative, convex, and non-decreasing function of $u \geqq 0$. If we use fundamental properties of convex functions, it is readily seen from (1.2) that

$$\int_0^{2\pi} \phi(|f'(x)|/n)\,dx \leqq \int_0^{2\pi} \frac{1}{n}\sum_{\nu=1}^{2n} \rho_\nu^{(n)} \phi(|f(x+\theta_\nu^{(n)})|)\,dx$$
$$= \int_0^{2\pi} \phi(|f(x)|)\,dx.$$

The result for L_p can be derived from this immediately by choosing $\phi(u) = u^p$, whence

$$\|f'\|_p \leqq n\,\|f\|_p. \tag{1.5}$$

This is a complete analogue of S. Bernstein's classical theorem which can be derived from (1.5) by allowing $p \to \infty$.

3. The corresponding problem for rational polynomials seems to be more complicated because no analogue of (1.2) is known in this case. We show, however, that A. Markoff's theorem still can be extended to the space L_p by proving the following

THEOREM. *Let $p \geqq 1$ and let $f(x)$ be an arbitrary rational polynomial of degree n. Then*

$$\left\{\int_{-1}^{+1} |f'(x)|^p\,dx\right\}^{1/p} : \left\{\int_{-1}^{+1} |f(x)|^p\,dx\right\}^{1/p} \leqq An^2, \tag{1.6}$$

where A is a constant which depends only on p, but not on $f(x)$ or on n.

For each n there exist polynomials $f(x)$ of degree n such that the left member of (1.6) *is $\geqq Bn^2$, where B is a constant of the same nature as A.*

A. Markoff's theorem (with a less precise value of the constant A) is obtained from (1.6) by allowing $p \to \infty$. Another important case, namely, $p = 2$, was treated some time ago by E. Schmidt.[3] The treatment of this special case given in Part V below is, as we understand, essentially identical with Schmidt's line of argument. In this special case the constant A above can be characterized in a more precise fashion than in the general case.

4. Neither customary methods used for the proof of A. Markoff's original theorem[4] ($p = \infty$) nor E. Schmidt's elegant method ($p = 2$) seems to be applicable in the general case. In Parts II and III we give two variants of our proof of (1.6) (the first of them valid only for $p > 1$). Both may present interest even in the limiting case $p = \infty$. In Part IV we show that the "order"

[3] *Die asymptotische Bestimmung des Maximums des Integrals über das Quadrat der Ableitung eines normierten Polynoms, dessen Grad ins Unendliche wächst*, Sitzungsberichte der Preussischen Akademie, 1932, p. 287. This note contains a statement of the result without proof.

[4] See, e.g., G. Pólya and G. Szegö, *Aufgaben und Lehrsätze aus der Analysis*, Berlin, 1925, vol. 2, pp. 91, 287, problem 23.

n^2 of the bound in (1.6) is "the best possible". The determination of the precise value of the constant A seems to be a more difficult problem.

II. General case. First method

1. Our first proof of the inequality (1.6) is based on the following two lemmas, one of which is due to Gabriel and the other is an extension of a classical theorem of S. Bernstein-M. Riesz.[5]

LEMMA 2.1. *If Γ is any convex closed curve in a complex z-plane and C any convex curve inside Γ, and if $F(z)$ is regular inside and on Γ, then*

$$\int_C |F(z)|^\lambda |dz| \leqq G \int_\Gamma |F(z)|^\lambda |dz|. \tag{2.1}$$

Here λ is any number $\geqq 0$ and G an absolute constant.[6]

Let C be any simple closed rectifiable Jordan curve in a complex z-plane and let

$$z = \psi(w) = cw + c_0 + c_1 w^{-1} + \cdots + c_n w^{-n} + \cdots, \; c > 0,$$

be the function which maps conformally the simply connected infinite domain exterior to C into the exterior of the unit circle $|w| = 1$ in the w-plane. Let C_R be the image in the z-plane of the circle $|w| = R$. With this notation we have

LEMMA 2.2. *If $f(z)$ is any polynomial of degree n, then*

$$\int_{C_R} |f(z)|^p |dz| \leqq R^{np+1} \int_C |f(z)|^p |dz|, \qquad p > 0. \tag{2.2}$$

Consider the function

$$\phi(w) = w^{-n} f(\psi(w))(\psi'(w))^{1/p}.$$

It is clear that this function is regular for $|w| > 1$ (including the point at infinity) and therefore the integral

$$I_R = \int_0^{2\pi} |\phi(Re^{i\theta})|^p \, d\theta \uparrow \int_0^{2\pi} |\phi(e^{i\theta})|^p \, d\theta$$

as $R \downarrow 1$.[7] Here $\phi(e^{i\theta}) = \lim_{R \to 1} \phi(Re^{i\theta})$. Now

$$\int_{C_R} |f(z)|^p |dz| = R^{np+1} \int_0^{2\pi} |\phi(Re^{i\theta})|^p \, d\theta = R^{np+1} I_R,$$

$$\int_C |f(z)|^p |dz| = \int_0^{2\pi} |\phi(e^{i\theta})|^p \, d\theta,$$

and (2.2) follows at once.

[5] Cf. M. Riesz, *Über einen Satz des Herrn Serge Bernstein*, Acta Mathematica, vol. 40 (1916), pp. 337–347.

[6] R. M. Gabriel, *Concerning integrals of moduli of regular functions along convex curves*, Proceedings of the London Math. Soc., (2), vol. 39 (1935), pp. 216–231, esp. p. 229.

[7] Concerning the monotony property of I_R cf. Pólya-Szegö, loc. cit., vol. 1, p. 144 and p. 330, problem 310. The fact that the limit function $\phi(e^{i\theta})$ exists and belongs to L_p over $(0, 2\pi)$ is well known and is trivial in the case which will occur in the subsequent discussion.

Remark. It is clear that the preceding argument can be applied in the case where C is a rectifiable Jordan arc. The only change will consist in replacing $\int_C$ by $2\int_C$. Thus we have

$$\int_{C_R} |f(z)|^p |dz| \leq 2R^{np+1} \int_C |f(z)|^p |dz|. \tag{2.3}$$

2. We now are prepared to give a proof of (1.6) in the case $p > 1$; this case will be assumed to hold throughout the remainder of this part. Let $f(x)$ be ar arbitrary polynomial of degree n. By Cauchy's formula we have

$$f'(x) = \frac{1}{2\pi i} \int_{C_x} \frac{f(z)}{(z - x)^2} dz; \tag{2.4}$$

here $-1 \leq x \leq +1$ and the contour of integration C_x which may depend on x will be specified later.[8] Hölder's inequality yields

$$\begin{aligned} \int_{-1}^{+1} |f'(x)|^p dx &\leq (2\pi)^{-p} \int_{-1}^{+1} dx \left| \int_{C_x} \frac{f(z)}{(z - x)^2} dz \right|^p \\ &\leq (2\pi)^{-p} \int_{-1}^{+1} dx \left\{ \int_{C_x} |f(z)|^p |dz| \right\} \left\{ \int_{C_x} \frac{|dz|}{|z - x|^{2p'}} \right\}^{p/p'}, \end{aligned} \tag{2.5}$$

where $1/p + 1/p' = 1$.

Let $R > 1$ be arbitrary and let E_R denote the ellipse with foci at ± 1 and semi-axes

$$a = \tfrac{1}{2}(R + R^{-1}), \qquad b = \tfrac{1}{2}(R - R^{-1}). \tag{2.6}$$

An elementary discussion[9] gives the following expression for the shortest distance $D = D(x, R)$ of x from E_R:

$$D(x, R) = \begin{cases} b(1 - x^2)^{\frac{1}{2}}, & \text{if } |x| \leq a^{-1}, \\ a - |x|, & \text{if } a^{-1} \leq |x| \leq 1. \end{cases} \tag{2.7}$$

Now choose for C_x the circle with the center x and radius D; this circle is internally tangent to E_R. Then the last factor in the right member of (2.5) does not exceed $(2\pi D^{1-2p'})^{p/p'} = (2\pi)^{p/p'} D^{-p-1}$. As for the first expression in braces in (2.5), a successive application of Lemma 2.1 and of the remark following Lemma 2.2 yields

$$\int_{C_x} |f(z)|^p |dz| \leq G \int_{E_R} |f(z)|^p |dz| \leq 2GR^{np+1} \int_{-1}^{+1} |f(x)|^p dx.$$

[8] An analogous argument, in the essentially simpler case $p = \infty$, was used by Montel, *Sur les polynomes d'approximation*, Bulletin de la Société Mathématique de France, vol. 46 (1918), pp. 151–192, esp. pp. 160–161.

[9] For $b^2u^2 + a^2v^2 = a^2b^2$ we have $\{(u - x)^2 + v^2\}^{\frac{1}{2}} = \{(a^{-1}u - ax)^2 + b^2(1 - x^2)\}^{\frac{1}{2}}$.

On substituting these results in (2.5) we have

$$\int_{-1}^{+1} |f'(x)|^p\,dx \leqq \pi^{-1}GR^{np+1}\int_{-1}^{+1}|f(x)|^p\,dx\cdot\int_{-1}^{+1}\{D(x,R)\}^{-p-1}\,dx. \tag{2.8}$$

3. To obtain the most favorable estimate for the right-hand member of (2.8) let $R \downarrow 1$, so that

$$a - 1 \cong \tfrac{1}{2}(R-1)^2, \qquad b \cong R - 1.$$

Then

$$\begin{aligned}\int_{-1}^{+1}\{D(x,R)\}^{-p-1}\,dx = 2b^{-p-1}\int_{0}^{1/a}(1-x^2)^{-\frac{1}{2}(p+1)}\,dx \\ + 2\int_{1/a}^{1}(a-x)^{-p-1}\,dx,\end{aligned} \tag{2.9}$$

where

$$\begin{aligned}\int_0^{1/a}(1-x^2)^{-\frac{1}{2}(p+1)}\,dx &\cong 2^{-\frac{1}{2}(p+1)}\int_0^{1/a}(1-x)^{-\frac{1}{2}(p+1)}\,dx \\ &\cong 2^{-\frac{1}{2}(p+1)}\frac{(1-1/a)^{-\frac{1}{2}(p-1)}}{\frac{1}{2}(p-1)} \cong \frac{(R-1)^{1-p}}{p-1},\end{aligned}$$

$$\begin{aligned}\int_{1/a}^{1}(a-x)^{-p-1}\,dx &= p^{-1}(a-1)^{-p}\{1-(1+1/a)^{-p}\} \\ &\cong 2^p p^{-1}(1-2^{-p})(R-1)^{-2p},\end{aligned}$$

so that finally

$$\int_{-1}^{+1}\{D(x,R)\}^{-p-1}\,dx = O\{(R-1)^{-2p}\}.$$

On putting $R = 1 + 1/n$ we obtain an upper bound for the right member in (2.9), which is $O\{(R-1)^{-2p}\} = O(n^{2p})$. On substituting into (2.8) we obtain inequality (1.6) in the case $p > 1$.

It should be observed that in the case $p = 1$ the same method yields an estimate $O(n^2\log n)$ for (2.9). This leads to a result much less precise than that obtained in Part III by the second method.

III. General case. Second method

1. Our second method in turn is based on two lemmas which may present an interest in themselves and on the corresponding result for trigonometric polynomials, which was stated in Part I.

Lemma 3.1. *If $p \geqq 1$ and $f(x)$ is an arbitrary polynomial ($\not\equiv 0$) of degree n, then*

$$\int_{-1}^{+1}|f(x)|^p(1-x^2)^{-\frac{1}{2}}\,dx < 2(np+1)^{np+1}(np)^{-np}\int_{-1}^{+1}|f(x)|^p\,dx. \tag{4.1}$$

LEMMA 3.2. *Under the assumptions of Lemma* 3.1 *we have*[10]

$$\int_{-1}^{+1} |f(x)|^p\,dx < \left(\frac{2}{p-1}\right)^{p-1}(np+p)^{np+p}(np+1)^{-np-1}\int_{-1}^{+1}|f(x)|^p(1-x^2)^{\frac{1}{2}(p-1)}\,dx. \tag{4.2}$$

In what follows we shall use the elementary transformation $x = \frac{1}{2}(w + w^{-1})$ which for $w = e^{i\theta}$ reduces to $x = \cos\theta$. We shall also use integrals extended over certain circles $|w| = \rho$. We therefore shall write $w = \rho e^{i\theta}$ and integrate with respect to θ over $0 \leqq \theta \leqq 2\pi$. Now, to prove Lemma 3.1, let $0 < r < 1$. Then[11]

$$\int_{|w|=r} |w^n f\{\tfrac{1}{2}(w+w^{-1})\}|^p \left|\frac{w^2-1}{4}\right| d\theta < \int_{|w|=1} |w^n f\{\tfrac{1}{2}(w+w^{-1})\}|^p \left|\frac{w^2-1}{4}\right| d\theta$$

$$= \tfrac{1}{2}\int_0^{2\pi} |f(\cos\theta)|^p\,|\sin\theta|\,d\theta = \int_{-1}^{+1} |f(x)|^p\,dx = I_1.$$

Since $|w^2 - 1| \geqq 1 - r^2$ for $|w| = r$, we have

$$\int_{|w|=r} |w^n f\{\tfrac{1}{2}(w+w^{-1})\}|^p\,d\theta < \frac{4}{1-r^2}I_1,$$

whence

$$\int_{-1}^{+1} |f(x)|^p(1-x^2)^{-\frac{1}{2}}\,dx = \tfrac{1}{2}\int_0^{2\pi}|f(\cos\theta)|^p\,d\theta$$

$$= \tfrac{1}{2}\int_{|w|=1} |w^n f\{\tfrac{1}{2}(w+w^{-1})\}|^p\,d\theta < \tfrac{1}{2}\int_{|w|=r^{-1}} |w^n f\{\tfrac{1}{2}(w+w^{-1})\}|^p\,d\theta$$

$$= \tfrac{1}{2}r^{-2np}\int_{|w|=r} |w^n f\{\tfrac{1}{2}(w+w^{-1})\}|^p\,d\theta < \frac{2r^{-2np}}{1-r^2}I_1.$$

Lemma 3.1 follows immediately if we put $1 - r^2 = (np+1)^{-1}$.

[10] The factor $\left(\frac{2}{p-1}\right)^{p-1}$ should be replaced by 1 when $p = 1$.

[11] To derive this inequality it is perhaps simplest to observe that if $F_1(z), \cdots, F_l(z)$ are any set of regular analytic functions, then

$$F(z) = |F_1(z)|^{p_1}\cdots|F_l(z)|^{p_l}, \qquad p_1 \geqq 0, \cdots, p_l \geqq 0,$$

is subharmonic, and to use well-known properties of subharmonic functions. Cf. Radó, *Subharmonic functions*, Ergebnisse der Mathematik, vol. 5, Berlin, 1937: see n. 8. 2.4.

A similar argument can be used in proving Lemma 3.2. Indeed, we have

$$\int_{-1}^{+1} |f(x)|^p dx = \tfrac{1}{2}\int_0^{2\pi} |f(\cos\theta)|^p\, |\sin\theta|\, d\theta$$

$$= \int_0^{2\pi} |w^n f\{\tfrac{1}{2}(w + w^{-1})\}|^p \left|\frac{w^2 - 1}{4}\right| d\theta < \int_{|w|=r^{-1}} |w^n f\{\tfrac{1}{2}(w + w^{-1})\}|^p \left|\frac{w^2 - 1}{4}\right| d\theta$$

$$= r^{-2np-2}\int_{|w|=r} |w^n f\{\tfrac{1}{2}(w + w^{-1})\}|^p \left|\frac{1 - w^2}{4}\right| d\theta$$

$$< \tfrac{1}{4} r^{-2np-2}(1 - r^2)^{1-p}\int_{|w|=r} |w^n f\{\tfrac{1}{2}(w + w^{-1})\}|^p\, |1 - w^2|^p\, d\theta$$

$$< \tfrac{1}{4} r^{-2np-2}(1 - r^2)^{1-p}\int_{|w|=1} |w^n f\{\tfrac{1}{2}(w + w^{-1})\}|^p\, |1 - w^2|^p\, d\theta$$

$$= 2^{p-1} r^{-2np-2}(1 - r^2)^{1-p} I_2,$$

where

$$I_2 = \int_{-1}^{+1} |f(x)|^p (1 - x^2)^{\frac{1}{2}(p-1)}\, dx.$$

Lemma 3.2 follows immediately if we write here

$$1 - r^2 = (p - 1)(np + p)^{-1}.$$

2. Inequality (1.6) now is readily derived by combining Lemmas 3.1 and 3.2 with Zygmund's result (1.5). Indeed, if we put again

$$I_1 = \int_{-1}^{+1} |f(x)|^p\, dx,$$

we have by Lemma 3.1

$$\int_0^{2\pi} |f(\cos\theta)|^p\, d\theta < 4(np + 1)^{np+1}(np)^{-np} I_1,$$

whence, by (1.5),

$$\int_0^{2\pi} |f'(\cos\theta)\sin\theta|^p\, d\theta < 4(np + 1)^{np+1}(np)^{-np} n^p I_1.$$

On replacing n by $(n - 1)$ in Lemma 3.2 we now have

$$\int_{-1}^{+1} |f'(x)|^p dx < \left(\frac{2}{p - 1}\right)^{p-1}(np)^{np}(np - p + 1)^{-np+p-1}\cdot 2(np + 1)^{np+1}(np)^{-np} n^p I_1$$

$$= 2^p(p - 1)^{1-p}(np + 1)^{np+1}(np - p + 1)^{-np+p-1} n^p I_1.$$

The coefficient of I_1 is $\cong (2ep)^p(p-1)^{1-p}n^{2p}$ as $n \to \infty$ for p fixed. This proves (1.6).

On the other hand, if we allow $p \to \infty$ for n fixed and observe that

$$\lim_{p\to\infty} 2(p-1)^{-1+1/p}(np+1)^{n+1/p}(np-p+1)^{-n+1-1/p} = 2\frac{n^{n+1}}{(n-1)^{n-1}} < 2en^2,$$

we obtain in the classical problem of A. Markoff the estimate

$$\|f'\| < 2en^2 \|f\|.$$

IV. Exactness of the order n^2

1. Let $f(x) = P_n^{(\alpha,\alpha)}(x)$ be Jacobi's polynomial in the "ultraspherical" case.[12] Then, as $n \to \infty$,

$$\int_{-1}^{+1} |f(x)|^p dx \cong n^{\alpha p-2}2^{\alpha p+1}\int_0^\infty |u^{-\alpha}J_\alpha(u)|^p u\,du, \tag{4.1}$$

provided that $(\alpha + \frac{1}{2})p > 2$. Here $J_\alpha(u)$ is Bessel's function of order α.

The special case $p = 1$ of this formula can be found in a paper by Szegö.[13] The following line of argument is slightly simpler than the one used there.

We use the formula of "Mehler's type"

$$\lim_{n\to\infty} n^{-\alpha}P_n^{(\alpha,\alpha)}\left(\cos\frac{u}{n}\right) = (u/2)^{-\alpha}J_\alpha(u), \tag{4.2}$$

which holds uniformly over every finite interval, and the estimate[14]

$$P_n^{(\alpha,\alpha)}(\cos\theta) = \theta^{-\alpha-\frac{1}{2}}O(n^{-\frac{1}{2}}), \qquad n^{-1} \leqq \theta \leqq \pi/2. \tag{4.3}$$

2. Let now ω be a fixed positive number. Then

$$\int_{-1}^{+1} |P_n^{(\alpha,\alpha)}(x)|^p dx = 2\int_0^{\pi/2} |P_n^{(\alpha,\alpha)}(\cos\theta)|^p \sin\theta\, d\theta = 2\int_0^{\omega/n} + 2\int_{\omega/n}^{\pi/2}.$$

The first term of the last sum according to (4.2) equals

$$\frac{2}{n}\int_0^\omega \left|P_n^{(\alpha,\alpha)}\left(\cos\frac{u}{n}\right)\right|^p \sin\frac{u}{n}\,du \cong n^{\alpha p-2}2^{\alpha p+1}\int_0^\omega |u^{-\alpha}J_\alpha(u)|^p u\,du.$$

The second term according to (4.3) is

$$O(1)\int_{\omega/n}^{\pi/2} \theta^{-(\alpha+\frac{1}{2})p}n^{-p/2}\theta\,d\theta = O(1)n^{\alpha p-2}\omega^{2-(\alpha+\frac{1}{2})p}.$$

Since ω can be taken arbitrarily large, this proves (4.1).

[12] Cf. the notation in Pólya-Szegö, loc. cit., vol. 2, pp. 93, 94, 292, 293, problem 98.

[13] *Asymptotische Entwicklungen der Jacobischen Polynome*, Schriften der Königsberger Gelehrten Gesellschaft, 1933, pp. 35–112, esp. p. 88.

[14] Cf. Szegö, loc. cit., pp. 74, 77.

3. Now we observe that[15]

$$f'(x) = \frac{d}{dx} P_n^{(\alpha,\alpha)}(x) = (\alpha + \tfrac{1}{2}(n+1)) P_{n-1}^{(\alpha+1,\alpha+1)}(x),$$

so that, by (4.1),

$$\int_{-1}^{+1} |f'(x)|^p \, dx \cong (n/2)^p n^{(\alpha+1)p-2} 2^{(\alpha+1)p+1} \int_0^\infty | u^{-\alpha-1} J_{\alpha+1}(u) |^p u \, du, \tag{4.4}$$

and the ratio of the integrals in the left members of (4.4) and (4.1) remains $> Bn^{2p}$, as stated in the theorem of Part I, §3.

V. E. Schmidt's case $p = 2$

1. Without loss of generality we can confine ourselves to real polynomials $f(x)$. Indeed, if we write $f(x) = g(x) + ih(x)$, where $g(x)$ and $h(x)$ have real coefficients, we have

$$\begin{aligned} \int_{-1}^{+1} |f(x)|^2 dx &= \int_{-1}^{+1} \{g(x)\}^2 dx + \int_{-1}^{+1} \{h(x)\}^2 dx, \\ \int_{-1}^{+1} |f'(x)|^2 dx &= \int_{-1}^{+1} \{g'(x)\}^2 dx + \int_{-1}^{+1} \{h'(x)\}^2 dx. \end{aligned} \tag{5.1}$$

Let now $f(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n$, $n \geqq 2$. Then the determination of the maximum of $\int_{-1}^{+1} \{f'(x)\}^2 \, dx$ under the condition $\int_{-1}^{+1} \{f(x)\}^2 dx = 1$ is a characteristic value problem leading to the system of equations

$$\frac{\partial}{\partial a_\nu}\left[\int_{-1}^{+1} \{f'(x)\}^2 dx - \lambda \int_{-1}^{+1} \{f(x)\}^2 dx\right] = 0 \quad (\nu = 0, 1, 2, \cdots, n). \tag{5.2}$$

This system is equivalent to the condition that

$$\int_{-1}^{+1} f'(x) q'(x) \, dx - \lambda \int_{-1}^{+1} f(x) q(x) \, dx = 0 \tag{5.3}$$

be satisfied by an arbitrary polynomial $q(x)$ of degree n. Integration by parts gives

$$\int_{-1}^{+1} \{ f''(x) + \lambda f(x) \} \, q(x) \, dx = f'(1) q(1) - f'(-1) q(-1). \tag{5.4}$$

2. Introducing Legendre polynomials $\{P_n(x)\}$ in the usual notation, we put in (5.3)

$$q(x) = \sum_{\nu=0}^{n} (\nu + \tfrac{1}{2}) P_\nu(x) P_\nu(y), \tag{5.5}$$

[15] Szegö, loc. cit., p. 38, (5).

where y is a parameter. By using familiar properties of Legendre polynomials, we obtain

$$
\begin{aligned}
(5.6)\quad f''(y) + \lambda f(y) &= f'(1) \sum_{\nu=0}^{n} (\nu + \tfrac{1}{2})P_\nu(y) - f'(-1) \sum_{\nu=0}^{n} (\nu + \tfrac{1}{2})P_\nu(-y) \\
&= \tfrac{1}{2}f'(1)\{P'_n(y) + P'_{n+1}(y)\} + \tfrac{1}{2}(-1)^n f'(-1)\{P'_n(y) - P'_{n+1}(y)\}.
\end{aligned}
$$

Now it is helpful to write

$$
(5.7)\qquad f(y) = \alpha_0 P'_{n+1}(y) + \alpha_1 P'_n(y) + \alpha_2 P'''_{n+1}(y) + \alpha_3 P'''_n(y) + \cdots.
$$

This is obviously possible if the real constants $\alpha_0, \alpha_1, \alpha_2, \cdots, \alpha_n$ are suitably chosen. On substituting (5.7) into (5.6) and comparing the coefficients we get

$$
\begin{aligned}
&\lambda\alpha_0 = \tfrac{1}{2}f'(1) - \tfrac{1}{2}(-1)^n f'(-1) = \alpha_0 P''_{n+1}(1) + \alpha_2 P^{(4)}_{n+1}(1) + \cdots, \\
(5.8)\quad &\lambda\alpha_1 = \tfrac{1}{2}f'(1) + \tfrac{1}{2}(-1)^n f'(-1) = \alpha_1 P''_n(1) + \alpha_3 P^{(4)}_n(1) + \cdots, \\
&\lambda\alpha_\nu + \alpha_{\nu-2} = 0 \qquad (\nu = 2, 3, \cdots, n).
\end{aligned}
$$

This system is equivalent to (5.2) and readily furnishes the characteristic values and functions. Indeed, since λ is real and positive, we have

$$
(5.9)\qquad \alpha_{2\nu} = (-\lambda)^{-\nu}\alpha_0, \qquad \alpha_{2\nu+1} = (-\lambda)^{-\nu}\alpha_1, \qquad (\nu \geqq 1),
$$

and

$$
(5.10)\qquad
\begin{aligned}
\lambda\alpha_0 &= \alpha_0\{P''_{n+1}(1) - \lambda^{-1}P^{(4)}_{n+1}(1) + \lambda^{-2}P^{(6)}_{n+1}(1) - \cdots\}, \\
\lambda\alpha_1 &= \alpha_1\{P''_n(1) - \lambda^{-1}P^{(4)}_n(1) + \lambda^{-2}P^{(6)}_n(1) - \cdots\}.
\end{aligned}
$$

Thus the set of characteristic values is obtained from the two equations

$$
(5.11)\qquad
\begin{aligned}
P_{n+1}(1) - \lambda^{-1}P''_{n+1}(1) + \lambda^{-2}P^{(4)}_{n+1}(1) - \cdots &= 0, \\
P_n(1) - \lambda^{-1}P''_n(1) + \lambda^{-2}P^{(4)}_n(1) - \cdots &= 0.
\end{aligned}
$$

3. Observe that we have for all ν and n

$$
(5.12)\qquad n^{-4\nu}P^{(2\nu)}_n(1) = n^{-4\nu}\frac{(n+2\nu)!}{(n-2\nu)!}\,\frac{1}{2^{2\nu}(2\nu)!} < \frac{2}{2^{2\nu}(2\nu)!},
$$

while for ν fixed

$$
(5.13)\qquad \lim_{n\to\infty} n^{-4\nu}P^{(2\nu)}_n(1) = \frac{1}{2^{2\nu}(2\nu)!}.
$$

Let $\lambda_{n0}, \lambda_{n1}, \cdots, \lambda_{nn}$ denote the characteristic values in decreasing order. If we observe (5.12), (5.13), and apply Hurwitz' theorem to each of the equations (5.11) we readily see that the limit

$$
(5.14)\qquad \lim_{n\to\infty} n^{-4}\lambda_{nk} = \lambda_k \qquad (k = 0, 1, 2, \cdots)
$$

exists for each fixed k and that furthermore the set $\{\lambda_k\}$ represents the set of all the roots of

$$\sum_{\nu=0}^{\infty} \frac{(-\lambda)^{-\nu}}{2^{2\nu}(2\nu)!} = \cos (2\lambda^{\frac{1}{2}})^{-1} = 0, \tag{5.15}$$

so that

$$\lambda_k = \pi^{-2}(2k+1)^{-2} \qquad (k = 0, 1, 2, \cdots). \tag{5.16}$$

If now M_n^2 denotes the maximum of the ratio of the integrals $\int_{-1}^{+1} \{f'(x)\}^2\, dx$ and $\int_{-1}^{+1} \{f(x)\}^2\, dx$, we have the final result[16]

$$\lim_{n\to\infty} n^{-2} M_n = \lim_{n\to\infty} n^{-2} \lambda_{n0}^{\frac{1}{2}} = \lambda_0^{\frac{1}{2}} = \pi^{-1}. \tag{5.17}$$

Remark. It can be shown that the largest characteristic value λ_{n0} necessarily is a root of the *first* of the equations (5.11). Indeed let $f_{n+1}(\lambda)$ and $f_n(\lambda)$ denote the left-hand members of the two equations (5.11), respectively. A simple calculation shows that

$$\frac{d}{d\lambda} \{\lambda^{\frac{1}{2}(n+1)}[f_{n+1}(\lambda) + f_n(\lambda)]\} = (n+1)\lambda^{\frac{1}{2}(n-1)} f_n(\lambda). \tag{5.18}$$

If $f_{n+1}(\lambda_{n0}) \neq 0$, we necessarily have $f_{n+1}(\lambda_{n0}) > 0$, $f'_{n+1}(\lambda_{n0}) > 0$, $f'_n(\lambda_{n0}) \geqq 0$, so that the left-hand member of (5.18) is positive for $\lambda = \lambda_{n0}$, whereas $f_n(\lambda_{n0}) = 0$. This is a contradiction.

In fact it is easy to show that λ_{n0} satisfies only the first of the two equations (5.11).

Yale University, Washington University, and Brown University.

[16] There is a slight discrepancy between this result and that of E. Schmidt according to which this limit would be 1 instead of π^{-1}.

$(n-1)$-DIMENSIONAL CHARACTERISTIC STRIPS OF A FIRST ORDER EQUATION AND CAUCHY'S PROBLEM

By E. W. Titt

Consider the first order partial differential equation

$$F(x^\alpha \mid z \mid p_\alpha) = 0 \qquad (p_\alpha = \partial z/\partial x^\alpha) \tag{a}$$

in one unknown z and n independent variables x^α ($n \geqq 3$). The purpose of the present paper is to generalize to the case of more than two independent variables the usual geometrical discussion of Cauchy's problem showing the manifolds for which the problem is indeterminate.[1] In doing this we introduce the concept of an $(n-1)$-dimensional characteristic strip and study its relation to the one-dimensional characteristic strips.

1. We first recall the geometrical approach to the one-dimensional characteristic strip. If we assume that the space S^{n+1} with coördinates $x^1, \cdots, x^n$, z is Euclidean with rectangular Cartesian coördinates, the problem of integrating the equation (a) is that of determining a hypersurface[2]

$$z = z(x^1, \cdots, x^n) \tag{1.1}$$

in S^{n+1} such that the direction ratios $p_1 : \cdots : p_n : -1$ of the normal to (1.1) satisfy the condition (a) at each point P of (1.1). The geometrical configuration consisting of a point $P(x^1, \cdots, x^n, z)$ and a hyperplane passing through P, namely,

$$Z - z = p_\alpha(X^\alpha - x^\alpha), \tag{1.2}$$

in called an element. In general, the integral elements at P, i.e., those which are possible tangent hyperplanes to integral hypersurfaces at P, envelope a hypercone T with vertex at P. It then follows easily that an integral element (1.2) is tangent to the hypercone T along the generator given by[3]

$$\frac{X^1 - x^1}{F_{p_1}} = \cdots = \frac{X^n - x^n}{F_{p_n}} = \frac{Z - z}{p_\alpha F_{p_\alpha}}. \tag{1.3}$$

Received June 16, 1937; developed in part while the author was a National Research Fellow and mentioned in a paper presented to the American Mathematical Society, February 23, 1935. See Abstract 41-3-103, Bull. Amer. Math. Soc., vol. 41 (1935), p. 182.

[1] See, for example, E. Goursat, *Cours d'Analyse Mathématique*, Tome II, 1924, Chapter 22. We do not discuss the regularity requirements on F or the initial manifold. For a treatment of this question in the case of two independent variables the reader is referred to G. A. Bliss, Princeton Colloquium Lectures, Amer. Math. Soc., 1913, p. 98.

[2] In S^{n+1} we shall call an n-dimensional spread, a hypersurface, and an $(n-1)$-dimensional spread, an edge. For linear spreads we shall use the terminology hyperplane and plane edge.

[3] Cf. Goursat, loc. cit., p. 616.

The characteristic curves are the curves traced on an integral hypersurface which are tangent at each point P to the generator (1.3) determined by the integral element at P. As is well known, these curves are determined by the system of ordinary equations

$$(1.4)\qquad \frac{dx^1}{F_{p_1}} = \cdots = \frac{dx^n}{F_{p_n}} = \frac{dz}{p_\alpha F_{p_\alpha}} = \frac{dp_1}{-F_{x^1} - F_z p_1} = \cdots = \frac{dp_n}{-F_{x^n} - F_z p_n} = dv^1$$

without knowledge of an integral hypersurface. The geometrical configuration consisting of a curve together with a strip of elements along it satisfying the system (1.4) will be called a one-dimensional characteristic strip.

2. Now let us turn to the consideration of Cauchy's problem for equation (a), i.e., the problem of passing an integral hypersurface (1.1) through an arbitrary edge

$$(2.1)\qquad x^\alpha = \xi^\alpha(v^1, \cdots, v^{n-1}); \qquad z = \zeta(v^1, \cdots, v^{n-1}).$$

Generalizing the idea of a characteristic curve, by a characteristic edge we shall mean an edge which lies on an integral hypersurface and which has the property that its tangent plane edge at each point P contains the generator (1.3) determined by the integral element at P. We shall now proceed to find for our characteristic edge a system of partial differential equations, which is independent of the integral hypersurface.

If the plane edge, tangent to the characteristic edge (2.1) at the point P, is to contain the generator (1.3) associated with P, then the rank of the matrix

$$(2.2)\qquad \left\| \begin{matrix} F_{p_1} & \cdots & F_{p_n} & p_\alpha F_{p_\alpha} \\ x^1_1 & \cdots & x^n_1 & z_1 \\ \cdots & \cdots & \cdots & \cdots \\ x^1_{n-1} & \cdots & x^n_{n-1} & z_{n-1} \end{matrix} \right\|,$$

where the subscripts on the x's and z denote partial derivatives with respect to the v's, must be less than n.[4] Since the characteristic edge (2.1) lies on an integral hypersurface, the integral element at P must contain the tangent plane edge at P, i.e.,

$$(2.3)\qquad z_a = p_\alpha x^\alpha_a.$$

Multiplying each of the first n columns of (2.2) by p_α and subtracting from the last, we have on account of (2.3) that the condition on (2.2) is completely equivalent to

$$(2.4)\qquad F_{p_\alpha} D_\alpha = 0,$$

[4] In what follows, when we use the term *edge* we always imply that the matrix obtained by striking out the first row and last column of (2.2) is of maximum rank. Throughout Greek letters $\alpha, \beta, \gamma, \cdots$ have the range $1, \cdots, n$; Latin letters $a, b, c, \cdots$, the range $1, \cdots, n-1$; and Latin letters i, j, k, the range $2, \cdots, n-1$.

where the D_α are the cofactors of F_{p_α} in the determinant of the first n columns of (2.2). Differentiating the equation (a) with respect to x^α we obtain

$$F_{p_\beta} p_{\alpha\beta} + F_z p_\alpha + F_{x^\alpha} = 0 \qquad \left(p_{\alpha\beta} = \frac{\partial^2 z}{\partial x^\alpha \, \partial x^\beta}\right). \tag{2.5}$$

If we adopt the notation A^a_α for the cofactor of the element x^α_a in the determinant of the first n columns of (2.2), then we have immediately the relations

$$F_{p_\alpha} D_\beta + x^\alpha_a A^a_\beta = 0. \tag{2.6}$$

If $\alpha = \beta$, then (2.6) is equivalent to (2.4); and the other equations (2.6) are obvious. We then multiply the equations,

$$\frac{\partial p_\alpha}{\partial v^a} = p_{\alpha\beta} x^\beta_a,$$

through by A^a_γ, sum on a, and make use of (2.5) and (2.6), getting

$$A^a_\beta \frac{\partial p_\alpha}{\partial v^a} - D_\beta(F_z p_\alpha + F_{x^\alpha}) = 0. \tag{2.7}$$

If in particular $D_1 \neq 0$, the equations (2.7) with $\beta = 1$ imply the remainder. For it follows from (2.6) that any vector $A^1_\beta, \cdots, A^{n-1}_\beta, D_\beta$ ($\beta \neq 1$) is either a zero vector or proportional to $A^1_1, \cdots, A^{n-1}_1, D_1$. Let us associate with each point of the characteristic edge (2.1) the integral element containing the tangent plane edge. Then any set of $(2n + 1)$ functions,

$$x^\alpha = \xi^\alpha(v), \qquad z = \zeta(v), \qquad p_\alpha = \pi_\alpha(v), \tag{2.8}$$

with $D_1 \neq 0$, which satisfy the system[5]

$$\begin{aligned} &\text{(a) } F_{p_\alpha} D_\alpha = 0, \qquad \text{(b) } z_a - p_\alpha x^\alpha_a = 0, \\ &\text{(c) } A^a_1 \frac{\partial p_\alpha}{\partial v^a} - D_1(F_z p_\alpha + F_{x^\alpha}) = 0, \end{aligned} \tag{2.9}$$

will be called an $(n - 1)$-*dimensional characteristic strip*.

We have immediately that *over an* $(n - 1)$-*dimensional characteristic strip the function* F = *constant*. For if we multiply the equation (2.9c) through by x^α_b, sum on α and make use of the equation (2.9b) we get

$$x^\alpha_a A^a_1 \frac{\partial p_\alpha}{\partial v^b} - D_1(F_z p_\alpha + F_{x^\alpha}) x^\alpha_b = 0.$$

This becomes

$$D_1 \frac{\partial F}{\partial v^b} = 0,$$

when use is made of (2.6) and (2.9b).

[5] In connection with (2.9a) let us notice that an edge lying on an integral hypersurface of a second order equation which satisfies the condition $F_{p_{\alpha\beta}} D_\alpha D_\beta = 0$ is the well-known characteristic surface. Let us also notice that in case $n = 2$ the system (2.9) reduces to (1.4).

3. Now let us show that *an $(n-1)$-dimensional characteristic strip is composed of an $(n-2)$-parameter family of one-dimensional characteristic strips.* In order to show that each one-dimensional characteristic strip determined by an element of our $(n-1)$-dimensional characteristic strip lies entirely on our $(n-1)$-dimensional strip, we consider the system of ordinary equations

$$\frac{dv^a}{d\bar{v}^1} = -\frac{A_1^a}{D_1}, \tag{3.1}$$

where the right members are functions of v^a alone in virtue of (2.8). By the existence theorem for ordinary equations the system (3.1) has a unique solution,

$$v^a = V^a(\bar{v}^1 \mid v_0^a), \tag{3.2}$$

which reduces to v_0^a for $\bar{v}^1 = 0$. The set of $(2n+1)$ functions of $\bar{v}^1$ obtained by substituting (3.2) into (2.8) constitutes a one-dimensional characteristic strip. For multiplying the equations (3.1) by x_a^α, summing on a, and making use of (2.9a) or (2.6) we find that the first n equations (1.4) are satisfied. The next equation (1.4) is a consequence of the first n and the equation (2.9b). The remaining equations (1.4) follow from (2.9c) and the equations (3.1).

Now let us consider any set of $(n-1)$ functions,

$$v_0^a = v_0^a(\bar{v}^2, \cdots, \bar{v}^{n-1}), \tag{3.3}$$

which for the set of values $\bar{v}^i = 0$ make the determinant

$$\begin{vmatrix} A_1^1[v_0^a(0)] & \cdots & A_1^{n-1}[v_0^a(0)] \\ \dfrac{\partial v_0^1(0)}{\partial \bar{v}^2} & \cdots & \dfrac{\partial v_0^{n-1}(0)}{\partial \bar{v}^2} \\ \cdots & \cdots & \cdots \\ \dfrac{\partial v_0^1(0)}{\partial \bar{v}^{n-1}} & \cdots & \dfrac{\partial v_0^{n-1}(0)}{\partial \bar{v}^{n-1}} \end{vmatrix} \neq 0. \tag{3.4}$$

On account of (3.4) the result of replacing the arbitrary constants v_0^a in (3.2) by (3.3), namely, $v^a = v^a(\bar{v}^b)$, can be regarded as a change of parameter in the neighborhood of the set of values $\bar{v}^a = 0$. In the transformed strip (2.8) the one-dimensional strips along which $\bar{v}^1$ varies are characteristic. Let the result of setting $\bar{v}^1 = 0$ in the transformed strip (2.8) be denoted by

$$x_0^\alpha(\bar{v}^2, \cdots, \bar{v}^{n-1}); \qquad z_0(\bar{v}^2, \cdots, \bar{v}^{n-1}); \qquad p_{\alpha 0}(\bar{v}^2, \cdots, \bar{v}^{n-1}). \tag{3.5}$$

In the neighborhood of the set of values $\bar{v}^i = 0$ the set of functions (3.5) satisfy the conditions

$$\begin{aligned} &\text{(a)}\ F[x_0^\alpha \mid z_0 \mid p_{\alpha 0}] = \text{const.}, \\ &\text{(b)}\ \frac{\partial z_0}{\partial \bar{v}^i} = p_{\alpha 0} \frac{\partial x_0^\alpha}{\partial \bar{v}^i}, \end{aligned} \tag{3.6}$$

and make the rank of the matrix

$$(3.7)\qquad \left\|\begin{matrix} F_{p_{10}} & \cdots & F_{p_{n0}} \\ \dfrac{\partial x_0^1}{\partial \bar{v}^2} & \cdots & \dfrac{\partial x_0^n}{\partial \bar{v}^2} \\ \cdots & \cdots & \cdots \\ \dfrac{\partial x_0^1}{\partial \bar{v}^{n-1}} & \cdots & \dfrac{\partial x_0^n}{\partial \bar{v}^{n-1}} \end{matrix}\right\|$$

equal to $(n - 1)$.

Conversely, let us show that *through any* $(n - 2)$*-dimensional strip* (3.5) *satisfying the conditions* (3.6) *and* (3.7) *there passes a unique* $(n - 1)$*-dimensional characteristic strip.* Suppose that we have an integral of the system (1.4) depending on an auxiliary variable $\bar{v}^1$ and $(2n + 1)$ arbitrary constants, x_0^α, z_0, $p_{\alpha 0}$ which reduces to x_0^α, z_0, $p_{\alpha 0}$ for $\bar{v}^1 = 0$. Let the result of substituting the $(2n + 1)$ functions (3.5) into this integral be denoted by S; we must show that S is an $(n - 1)$-dimensional characteristic strip. S satisfies the equation (2.9a) as a direct consequence of the first n equations (1.4). Also on account of the same equations (1.4) the quantities A_α^a which do not vanish identically will be A_α^1. In particular, let us consider $A_1^1 \neq 0$. Since $A_1^1 = -D_1$, the equations (2.9c) are identical with the last n equations (1.4). When $a = 1$, the equation (2.9b) follows from the first $(n + 1)$ equations (1.4). In order to show that the remaining equations (2.9b) are satisfied we put

$$V_i \equiv z_i - p_\alpha x_i^\alpha .$$

Differentiate this relation with respect to $\bar{v}^1$, make use of (2.9b) for $a = 1$ and equations (1.4), and obtain

$$(3.8)\qquad \frac{\partial V_i}{\partial \bar{v}^1} = F_{p_\alpha} \frac{\partial p_\alpha}{\partial \bar{v}^i} + F_z p_\alpha x_i^\alpha + F_{x^\alpha} x_i^\alpha .$$

From (3.6a) we have $\partial F / \partial \bar{v}^i = 0$, and equations (3.8) become

$$\frac{\partial V_i}{\partial \bar{v}^1} = -F_z V_i .$$

Since $V_i = 0$ for $\bar{v}^1 = 0$ by (3.6b), we have $V_i \equiv 0$.

The proof that the strip S is unique consists in showing that any other $(n - 1)$-dimensional characteristic strip S^* containing the $(n - 2)$-dimensional strip (3.5) must contain all of the $(n - 2)$-parameter family of one-dimensional characteristic strips which constitute S. By assumption, the initial element of any one of the one-dimensional characteristic strips in S lies also on S^*. By the argument in the first part of §3, the strip S^* consists of an $(n - 2)$-parameter family of one-dimensional characteristic strips. Since a one-dimensional characteristic strip is determined by its initial element, the argument is complete.

The construction of our $(n - 1)$-dimensional characteristic strip can be given

a geometrical interpretation. Let an initial $(n-2)$-dimensional manifold m be given by the first $(n+1)$ equations (3.5). Suppose that there exists an element E which contains the $(n-2)$-spread l tangent to m at P ($\bar{v}^i = 0$) and which is tangent to the hypercone T with vertex at P. Suppose also that l does not contain the generator through P determined by E. Then the equations (3.6) determine a one-parameter family of strips (3.5) containing the manifold m. Then as E describes one of these strips containing m, the $(n-2)$-parameter family of one-dimensional characteristic strips thus determined constitutes the $(n-1)$-dimensional characteristic strip.

4. Let us recall that Cauchy's method of integrating the equation (a) is to replace the arbitrary constants in a solution of (1.4) by $(2n+1)$ functions of $(n-1)$ parameters, namely,

$$
\begin{aligned}
&x_0^\alpha = \bar{\xi}^\alpha(v^2, \cdots, v^n); \qquad z_0 = \bar{\zeta}(v^2, \cdots, v^n); \\
&p_{\alpha 0} = \bar{\pi}_\alpha(v^2, \cdots, v^n),
\end{aligned}
\tag{4.1}
$$

satisfying the conditions

$$
\begin{aligned}
&F[x_0^\alpha \mid z_0 \mid p_{\alpha 0}] = 0, \\
&\frac{\partial z_0}{\partial v^\sigma} = p_{\alpha 0} \frac{\partial x_0^\alpha}{\partial v^\sigma}, \qquad (\sigma = 2, \cdots, n).
\end{aligned}
\tag{4.2}
$$

We now make a few remarks to show how the above theory can be applied in a discussion of Cauchy's problem. Let the initial edge M be given by the first $(n+1)$ equations (4.1) and let us suppose that there exists a set of values $p_{\alpha 0}$ which satisfy the system (4.2) at some point $P(v_0^\sigma)$ of M. First consider the case in which

$$
F_{p_\alpha} D_\alpha \neq 0 \tag{4.3}
$$

for the set of values $p_{\alpha 0}$, $\bar{\xi}^\alpha(v_0^\sigma)$, $\bar{\zeta}(v_0^\sigma)$. Hence we can solve equations (4.2) for the quantities $p_{\alpha 0}$ as functions of $v^2, \cdots, v^n$. On account of (4.3) the strip thus obtained will give us a solution of Cauchy's problem in the form (1.1). Geometrically we have assumed that there exists an element E, which contains the plane edge L tangent to M at P, and which is tangent to the hypercone T with vertex at P. We have supposed also that L does not contain the generator through P determined by E. Then the $(n-1)$-parameter family of one-dimensional characteristic strips, determined by E as P describes a neighborhood of itself on M, comprise our integral hypersurface. This integral hypersurface can also be thought of as generated by the one-parameter family of $(n-1)$-dimensional characteristic strips determined by any one-parameter family of $(n-2)$-dimensional spreads lying on the edge M.

Next let us consider the case where M bears an $(n-1)$-dimensional characteristic strip (4.1) and Cauchy's problem becomes indeterminate. Select any $(n-2)$-dimensional spread m lying on M such that at any point P of m the

$(n - 2)$-spread l tangent to m at P does not contain the generator through P determined by E. Through P on m pass a line t, on E, which with l determines a plane edge that does not contain the generator determined by E. Let Γ be a curve through P with t for its tangent line. Select a differentiable $(n - 2)$-parameter family of such curves Γ, one through each point of m, which taken together comprise an edge $\mathfrak{M}$. The plane edge tangent to $\mathfrak{M}$ at any point P of m lies in E but does not contain the generator through P determined by E. Therefore the integral hypersurface passing through $\mathfrak{M}$ must contain the $(n - 1)$-dimensional characteristic strip through m, i.e., the strip (4.1).

Any edge M lying on an integral hypersurface, whose tangent plane edge L at each point P contains the generator through P determined by E but which does not bear an $(n - 1)$-dimensional characteristic strip, must contain singularities of the integral hypersurface. For if the integral hypersurface possessed second derivatives at each point of M, we should conclude by the argument in §2 that M bears an $(n - 1)$-dimensional characteristic strip.

PRINCETON UNIVERSITY.

THEOREMS ON FOURIER SERIES AND POWER SERIES

By H. R. Pitt

1. Introduction

1.1. **Notation.** If $p > 0$, we write

$$\mathfrak{S}_p[a_n] = \left(\sum_{-\infty}^{\infty} |a_n|^p\right)^{1/p},$$

$$\mathfrak{T}_p[c_n] = \left(\int_0^1 \left(\sum_0^{\infty} |c_n| x^n\right)^p dx\right)^{1/p},$$

$$\mathfrak{M}_p[F(\theta)] = \left(\int_{-\pi}^{\pi} |F(\theta)|^p d\theta\right)^{1/p}.$$

If $g(z)$ is a regular analytic function for $|z| < 1$, we write

$$\mathfrak{H}_p[g(z)] = \lim_{r\to 1}\left(\int_{-\pi}^{\pi} |g(re^{i\theta})|^p d\theta\right)^{1/p}.$$

(The limit exists, since the expression in the bracket increases with r.)

1.2. Suppose that $F(\theta)$ is periodic and integrable and that $g(z)$ is regular in $|z| < 1$. Let

$$F(\theta) \sim \sum_{-\infty}^{\infty} a_n e^{ni\theta} \qquad (a_0 = 0), \tag{1.2.1}$$

$$g(z) = \sum_1^{\infty} c_n z^n \qquad (|z| < 1). \tag{1.2.2}$$

We shall prove that if p, q, α satisfy certain conditions,

$$\mathfrak{S}_q[c_n n^{-\lambda}] \leqq K\mathfrak{H}_p[g(z)(1 - z)^{\alpha}], \tag{1.2.3}$$

$$\mathfrak{S}_q[a_n n^{-\lambda}] \leqq K\mathfrak{M}_p[F(\theta)\theta^{\alpha}], \tag{1.2.4}$$

where

$$\lambda = \frac{1}{p} + \frac{1}{q} + \alpha - 1, \tag{1.2.5}$$

and the constants $K(p, q, \alpha)$ are independent of $g(z)$ and $F(\theta)$.

Special cases of these inequalities, due to Hausdorff[1] and Hardy and Little-

Received October 16, 1937.

[1] Hausdorff [5], Theorem II. (Numbers in brackets refer to the references at the end of the paper.) This is the case $\alpha = \gamma = 0$ of (1.2.4).

wood,[2] are well known. They can be derived from Theorems 1 and 2 of this paper by substituting appropriate values of p, q and α.

2. A theorem for power series

2.1. Our principal result is as follows.

THEOREM 1. *Let*

$$g(z) = \sum_{n=1}^{\infty} c_n z^n \qquad (|z| < 1);$$

$$\infty > q \geqq p > 0, \qquad \alpha \geqq 0,$$

$$\lambda = \frac{1}{p} + \frac{1}{q} + \alpha - 1 \geqq 0.$$

Then

(A) $$\mathfrak{S}_q[c_n n^{-\lambda}] \leqq K \mathfrak{H}_p[g(z)(1 - z)^{\alpha}],$$

(B) $$\mathfrak{T}_q[c_n n^{-\lambda-1+2/q}] \leqq K \mathfrak{H}_p[g(z)(1 - z)^{\alpha}],$$

whenever the right side is finite.

We shall denote the inequalities (A) and (B), for particular values of p, q and α, by $A[p, q, \alpha]$ and $B[p, q, \alpha]$, and use the symbol $\supset$ to show relations of inclusion between them. For example, we write

$$A[p, q, \alpha] \supset B[r, s, \beta],$$

if $B[r, s, \beta]$ can be deduced from $A[p, q, \alpha]$.

2.2. LEMMA 1.

(a) *Let*

$$\infty \geqq q \geqq 1, \qquad a < 1, \qquad q'b < 1;$$

$$c_n = n^{a+b-1} \sum_{\nu=1}^{n-1} \frac{b_{n-\nu}}{\nu^a (n - \nu)^b}.$$

Then

$$\mathfrak{S}_q[c_n] \leqq K(a, b, q) \mathfrak{S}_q[b_n].$$

(b) *Let*

$$r > 1, \qquad s > 1, \qquad a < \frac{1}{s'}, \qquad a \leqq \frac{2}{s'} - \frac{1}{r'},$$

$$0 \leqq 2a + c = \frac{2}{s'} - \frac{1}{r'} \leqq 1;$$

$$c_n = n^{-c} \sum_{\nu=1}^{n-1} \frac{b_\nu b_{n-\nu}}{\nu^a (n - \nu)^a}.$$

[2] Hardy and Littlewood [1] and [4]. The former covers the cases $\alpha = 0$ or $\gamma = 0$, while the latter deals with the case $\alpha = 1/p'$, $\lambda = 1/q$ of (1.2.3).

Then

$$\mathfrak{S}_r[c_n] \leqq K(a, r, s)\mathfrak{S}_s^2[b_n].$$

These results are special cases of a very general inequality of Hardy and Littlewood, Theorem 1 of [2]. We obtain (a) (with b_n', c_n' instead of b_n, c_n) by writing

$$p = \infty, \qquad r = q, \qquad \alpha = a, \qquad \beta = b + \frac{1}{q}, \qquad \gamma = a + b + \frac{1}{q} - 1,$$

$$a_n = n^{-a}, \qquad b_n = b_n' n^{-b}, \qquad c_n = c_n' n^{-a-b+1}.$$

We obtain (b) by writing

$$p = q = s, \qquad \alpha = \beta = a + \frac{1}{s}, \qquad \gamma = \frac{1}{r} - c,$$

$$a_n = b_n = b_n' n^{-a}, \qquad c_n = c_n' n^{c}.$$

2.3. Lemma 2.
(a) $A[p, q, 0]$ *is true if*

$$\infty > q \geqq p > 1, \qquad \frac{1}{p} + \frac{1}{q} \geqq 1.$$

(b) $A\left[p, q, 1 - \frac{1}{p} - \frac{1}{q}\right]$ *is true if*

$$\infty > q \geqq p > 1, \qquad \frac{1}{p} + \frac{1}{q} \leqq 1.$$

These results follow at once from Theorems 9 and 10 of [1]. In fact, the latter results are true for general Fourier series, whereas we are concerned here only with Fourier power series.

2.4. Lemma 3.

$$A[p, q, \alpha] \supset B[p, q, \alpha] \qquad \text{if } q > 1;$$

$$B[p, q, \alpha] \supset A[p, q, \alpha] \qquad \text{if } q \leqq 1.$$

These are immediate consequences of the following inequalities of Hardy and Littlewood.[3]

$$\mathfrak{T}_q[c_n] \leqq K(q)\mathfrak{S}_q[c_n n^{1-2/q}] \qquad \text{if } q > 1;$$

$$\mathfrak{S}_q[(c_n n^{1-2/q}] \leqq K(q)\mathfrak{T}_q[c_n] \qquad \text{if } q \leqq 1.$$

2.5. Lemma 4. *If* $\infty \geqq q \geqq 1$, $\infty > p > 0$, $\epsilon > 0$, $\alpha \geqq 0$, *then*

$$A[p, q, \alpha] \supset A[p, q, \alpha + \epsilon].$$

[3] Hardy and Littlewood [3], Theorems 3 and 11. (There is a misprint in Theorem 11. The exponent $(p + q - pq)/q$ on the left of (4.11) should be negative.) The case $q = 1$ is trivial.

Let

$$\phi(z) = \sum_1^\infty n^{\epsilon-1} z^n \qquad (|z| < 1).$$

Then[4] $\phi(z)$ is regular for $|z| \leqq 1$, $z \neq 1$, and has no zero, except at $z = 0$, in $|z| \leqq 1$. Moreover, $|\phi(z)(1-z)^\epsilon|^{-1}$ is bounded in $|z| \leqq 1$. Hence, if we write

$$z\omega(z) = g(z)/\phi(z) = \sum_1^\infty b_n z^n \qquad (|z| < 1),$$

we have

$$c_n = \sum_{\nu=1}^{n-1} \frac{b_{n-\nu}}{\nu^{1-\epsilon}},$$

$$\mathfrak{H}_p[\omega(z)z(1-z)^\alpha] \leqq K(\epsilon)\mathfrak{H}_p[g(z)(1-z)^{\alpha+\epsilon}].$$

We can write

$$c_n n^{-\lambda-\epsilon} = n^{-\lambda-\epsilon} \sum_{\nu=1}^{n-1} \frac{b_{n-\nu}(n-\nu)^{-\lambda}}{\nu^{1-\epsilon}(n-\nu)^{-\lambda}},$$

and since

$$1 - \epsilon < 1, \qquad -\lambda = \frac{1}{q'} - \alpha - \frac{1}{p} < \frac{1}{q'},$$

it follows from Lemma 1(a) that

$$\mathfrak{S}_q[c_n n^{-\lambda-\epsilon}] \leqq K\mathfrak{S}_q[b_n n^{-\lambda}].$$

Hence, assuming $A[p, q, \alpha]$, we have

$$\mathfrak{S}_q[c_n n^{-\lambda-\epsilon}] \leqq K\mathfrak{S}_q[b_n n^{-\lambda}] \leqq K\mathfrak{H}_p[\omega(z)z(1-z)^\alpha] \leqq K\mathfrak{H}_p[g(z)(1-z)^{\alpha+\epsilon}].$$

This gives $A[p, q, \alpha + \epsilon]$.

2.6. Lemma 5. *Let*

$$(2.6.1) \qquad q > 1, \qquad s > 1, \qquad s \geqq p > 0, \qquad \alpha \geqq 0,$$

$$\frac{2}{s} - 1 \leqq \frac{1}{q} \leqq \frac{2}{s}.$$

Then

$$A[p, s, \alpha] \supset A[\tfrac{1}{2}p, q, 2\alpha].$$

By using a theorem of F. Riesz,[5] we may suppose that $g(z)$ has no zeros in $|z| < 1$, except at $z = 0$, and write

$$zg(z) = [\omega(z)]^2, \qquad \omega(z) = \sum_1^\infty b_\nu z^\nu,$$

[4] See Hardy and Littlewood [4], page 367.

[5] F. Riesz [6]. See also page 207 of [1].

so that

$$c_{n-1} = \sum_{\nu=1}^{n-1} b_\nu b_{n-\nu}.$$

We suppose that $A[p, s, \alpha]$ is true. Then

$$\mathfrak{S}_s[b_n n^{-\mu}] \leqq K\mathfrak{H}_p[\omega(z)(1-z)^\alpha], \tag{2.6.2}$$

where

$$\mu = \frac{1}{p} + \frac{1}{s} + \alpha - 1.$$

Let

$$\lambda = \frac{2}{p} + \frac{1}{q} + 2\alpha - 1.$$

Then

$$c_{n-1}n^{-\lambda} = n^{-\lambda} \sum_{\nu=1}^{n-1} \frac{b_\nu \nu^{-\mu} b_{n-\nu}(n-\nu)^{-\mu}}{\nu^{-\mu}(n-\nu)^{-\mu}}.$$

It is plain from (2.6.1) that

$$-\mu < \frac{1}{s'}, \qquad -\mu \leqq \frac{2}{s'} - \frac{1}{q'}, \qquad 0 \leqq \lambda - 2\mu = \frac{2}{s'} - \frac{1}{q'} \leqq 1.$$

It follows from Lemma 1(b) that

$$\mathfrak{S}_q[c_{n-1}n^{-\lambda}] \leqq K\mathfrak{S}_s^2[b_n n^{-\mu}] \leqq K\mathfrak{H}_p^2[\omega(z)(1-z)^\alpha],$$

by (2.6.2). Hence

$$\mathfrak{S}_q[c_{n-1}n^{-\lambda}] \leqq K\mathfrak{H}_{\frac{1}{2}p}[\omega(z)^2(1-z)^{2\alpha}] = K\mathfrak{H}_{\frac{1}{2}p}[g(z)(1-z)^{2\alpha}],$$

which is equivalent to $A[\frac{1}{2}p, q, 2\alpha]$.

2.7. **Lemma 6.** *If* $\infty > q \geqq p > 0$, $\alpha \geqq 0$, *then*

$$B[p, q, \alpha] \supset B[\tfrac{1}{2}p, \tfrac{1}{2}q, 2\alpha].$$

As in Lemma 5, we may[6] suppose that $g(z) \neq 0$ for $|z| < 1$, $z \neq 0$, and write

$$zg(z) = [\omega(z)]^2, \qquad \omega(z) = \sum_{\nu=1}^{\infty} b_\nu z^\nu, \qquad c_{n-1} = \sum_{\nu=1}^{n-1} b_\nu b_{n-\nu}.$$

We suppose that $B[p, q, \alpha]$ is true, so that

$$\mathfrak{T}_q[b_n n^{-\mu-1+2/q}] \leqq K\mathfrak{H}_p[\omega(z)(1-z)^\alpha],$$

where

$$\mu = \frac{1}{p} + \frac{1}{q} + \alpha - 1.$$

[6] We use the fact that $\mathfrak{M}_r[\phi + \psi] \leqq K(r)[\mathfrak{M}_r(\phi) + \mathfrak{M}_r(\psi)]$ for any functions ϕ, ψ and any positive r. If $r \geqq 1$, this is Minkowski's inequality. If $r < 1$, it follows from the inequalities $a + b \leqq (a^r + b^r)^{1/r} \leqq 2^{1/r}(a + b)$.

Let

$$\lambda = \frac{2}{p} + \frac{2}{q} + 2\alpha - 1.$$

Then

$$\sum_{n=1}^{\infty} |c_{n-1}| n^{-\lambda-1+4/q} x^n \leqq \sum_{n=1}^{\infty} n^{-\lambda-1+4/q} \sum_{\nu=1}^{n-1} |b_\nu| |b_{n-\nu}|$$
$$= \sum_{n=1}^{\infty}\sum_{\nu=1}^{\infty} |b_\nu| x^\nu |b_n| x^n (n+\nu)^{-\lambda-1+4/q}$$
$$\leqq \left(\sum_{\nu=1}^{\infty} |b_\nu| x^\nu \nu^{-\mu-1+2/q}\right)^2,$$

since

$$\lambda + 1 - 4/q = 2(\mu + 1 - 2/q) \geqq 0$$

and $(a+b)^2 \geqq ab$ for any positive a, b. Hence

$$\mathfrak{T}_{\frac{1}{2}q}[c_{n-1} n^{-\lambda-1+4/q}] \leqq \mathfrak{T}_q^2[b_n n^{-\mu-1+2/q}] \leqq K\mathfrak{H}_p^2[\omega(z)(1-z)^\alpha]$$
$$= K\mathfrak{H}_{\frac{1}{2}p}[\omega(z)^2(1-z)^{2\alpha}] = K\mathfrak{H}_{\frac{1}{2}p}[g(z)z(1-z)^{2\alpha}],$$

and this is equivalent to $B[\frac{1}{2}p, \frac{1}{2}q, 2\alpha]$.

2.8. **Proof of Theorem 1.** We suppose first that $\infty > q \geqq p > 1$. In view of Lemma 3, it is sufficient to prove $A[p, q, \alpha]$, and this follows at once from Lemmas 2 and 4.

Next, let $\infty > q > 1 \geqq p > 0$. Because of Lemmas 3 and 4, it is sufficient to prove $A[p, q, 0]$. We suppose first that $q > 1 \geqq p > \frac{1}{2}$. We can choose s so that

$$2 - \frac{1}{p} \leqq \frac{2}{s} \leqq \frac{1}{q} + 1, \qquad \frac{1}{q} \leqq \frac{2}{s} \leqq \frac{1}{p}.$$

Then

$$\frac{2}{s} - 1 \leqq \frac{1}{q} \leqq \frac{2}{s}, \qquad s \geqq 2p > 1,$$

so that

$$A[2p, s, 0] \supset A[p, q, 0],$$

by Lemma 5. Moreover, since

$$\infty > s \geqq 2p > 1, \qquad \frac{1}{s} + \frac{1}{2p} \geqq 1,$$

it follows from Lemma 2(a) that $A[2p, s, 0]$ is true. Hence $A[p, q, 0]$ is true for $q > 1 \geqq p > \frac{1}{2}$. If we now put $s = q$ in Lemma 5, we can prove that $A[p, q, 0]$ is true for

$$\frac{1}{2} \geqq p > \frac{1}{4}, \cdots, \frac{1}{2^k} \geqq p > \frac{1}{2^{k+1}},$$

successively, and it follows that the result is true generally for $q > 1 \geqq p > 0$.

Finally, we have to consider the case $1 \geqq q \geqq p > 0$. Because of Lemma 3, it is sufficient to prove $B[p, q, \alpha]$. We choose an integer k so that $2 \geqq 2^k q > 1$. Then $B[2^k p, 2^k q, \alpha]$ is true for $\alpha \geqq 0$, by what we have already proved, and the conclusion follows by repeated application of Lemma 6. This completes the proof of Theorem 1.

2.9. We can express Theorem 1 in a slightly different form. We know[7] that if $\mathfrak{H}_p[g(z)(1 - z)^\alpha]$ is finite, then $g(z)$ has a boundary function $G(\theta)$ such that the ratio

$$\mathfrak{H}_p[g(z)(1 - z)^\alpha] : \mathfrak{M}_p[G(\theta)\theta^\alpha]$$

lies between positive bounds $K(p, \alpha)$. Conversely, if $\mathfrak{M}_p[G(\theta)\theta^\alpha]$ is finite, then $G(\theta)$ is the boundary function of an analytic function $g(z)$, and the same relation holds. It follows that we may replace $\mathfrak{H}_p[g(z)(1 - z)^\alpha]$ by $\mathfrak{M}_p[G(\theta)\theta^\alpha]$ in the conclusion of Theorem 1.

3. Theorems for Fourier series

3.1. We shall now show that the inequality (1.2.4) can be deduced from Theorem 1 when p, q and α satisfy certain extra conditions.

THEOREM 2. *Suppose that $F(\theta)$ is integrable and periodic. Let*

$$F(\theta) \sim \sum_{-\infty}^{\infty} a_n e^{ni\theta}, \qquad a_0 = 0;$$

$$\infty > q \geqq p > 1, \qquad \frac{1}{p'} > \alpha \geqq 0,$$

$$\lambda = \frac{1}{p} + \frac{1}{q} + \alpha - 1 \geqq 0.$$

Then

(a) $$\mathfrak{S}_q[a_n n^{-\lambda}] \leqq K \mathfrak{M}_p[F(\theta)\theta^\alpha],$$

(b) $$\mathfrak{M}_q[F(\theta)\theta^{-\lambda}] \leqq K \mathfrak{S}_p[a_n n^\alpha].$$

Let $G(\theta)$ be the conjugate of $F(\theta)$. Then $F(\theta) + iG(\theta)$ and $F(\theta) - iG(\theta)$ are boundary functions of $2\sum_1^\infty a_n z^n$ and $2\sum_1^\infty a_{-n} z^n$, respectively, and the conclusion (a) follows from Theorem 1 and the remarks of 2.9 if we show that

$$\mathfrak{M}_p[G(\theta)\theta^\alpha] \leqq K \mathfrak{M}_p[F(\theta)\theta^\alpha].$$

This has been proved by Hardy and Littlewood[8] when $p > 1$, $-1/p < \alpha < 1/p'$, and these conditions are plainly satisfied here.

The inequality (b) can be deduced from (a) by a "conjugacy" argument. Let

$$Q(\theta) = \sum c_n e^{ni\theta}$$

[7] F. Riesz [6], Theorems II and III. See also §4.1 of [4].

[8] Hardy and Littlewood [4], §6.3.

be a polynomial. Then

$$\int_{-\pi}^{\pi} F(\theta)\,\overline{Q(\theta)}\,d\theta = \sum a_n \bar{c}_n,$$

$$\left| \int_{-\pi}^{\pi} F(\theta)\theta^{-\lambda}\,\overline{Q(\theta)}\,\theta^{\lambda}\,d\theta \right| = \left| \sum a_n n^{\alpha} \bar{c}_n n^{-\alpha} \right|$$
$$\leqq \mathfrak{S}_p[a_n n^{\alpha}]\,\mathfrak{S}_{p'}[c_n n^{-\alpha}]$$
$$\leqq \mathfrak{S}_p[a_n n^{\alpha}] K \mathfrak{M}_{q'}[Q(\theta)\,\theta^{\lambda}]$$

by Theorem 2(a), since

$$\alpha = \frac{1}{p'} + \frac{1}{q'} + \lambda - 1 \geqq 0, \qquad \frac{1}{q} > \lambda \geqq 0, \qquad \infty > p' \geqq q' > 1.$$

It follows from the converse of Hölder's inequality that

$$\mathfrak{M}_q[F(\theta)\theta^{-\lambda}] \leqq K\mathfrak{S}_p[a_n n^{\alpha}].$$

This is what we require.

3.2. We can deduce from Theorem 1 the following extensions of Theorems 9 and 10 of [4].

THEOREM 3. *Suppose that* $G(\theta)$ *is the boundary function of an analytic function* $g(z) = \sum_1^{\infty} c_n z^n$, *that*

$$s_n(x) = \sum_1^n c_\nu e^{\nu i x};$$

and that p, q, α *satisfy the conditions of Theorem* 1. *Then*

$$\mathfrak{S}_q[(s_n(x) - s(x))n^{-\lambda}] \leqq K\mathfrak{M}_p[(F(x+\theta) - s(x))\theta^{\alpha-1}],$$

whenever the right side is finite.

THEOREM 4. *Suppose that* $F(\theta)$ *is integrable and*

$$F(\theta) \sim \sum_{-\infty}^{\infty} a_n e^{ni\theta}, \qquad a_0 = 0.$$

Let

$$s_n(x) = \sum_{-n}^{n} c_\nu e^{\nu i x},$$

$$\phi(x, \theta) = \tfrac{1}{2}[F(x+\theta) + F(x-\theta) - 2s(x)];$$

$$\infty > q \geqq p > 1, \qquad 1 + \frac{1}{p'} > \alpha \geqq 0,$$

$$\lambda = \frac{1}{p} + \frac{1}{q} + \alpha - 1 \geqq 0.$$

Then

$$\mathfrak{S}_q[(s_n(x) - s(x))n^{-\lambda}] \leqq K\mathfrak{M}_p[\phi(x, \theta)\theta^{\alpha-1}],$$

whenever the right side is finite.

We obtain Theorem 3 immediately on applying Theorem 1 to the function $(g(z) - s(0))(1 - z)^{-1}$. Theorem 4 follows from Theorem 1 by the argument given in §5.2 of [4].

The Hardy-Littlewood theorems are given, in each case, by $\alpha = 1/p'$.

References

1. G. H. Hardy and J. E. Littlewood, *Some new properties of Fourier constants*, Math. Ann., vol. 97 (1926), pp. 159–209.
2. G. H. Hardy and J. E. Littlewood, *An inequality*, Math. Zeitschr., vol. 40 (1935), pp. 1–40.
3. G. H. Hardy and J. E. Littlewood, *Elementary theorems concerning power series with positive terms*, Journ. für Math., vol. 157 (1927), pp. 141–158.
4. G. H. Hardy and J. E. Littlewood, *Some more theorems concerning Fourier series and Fourier power series*, this Journal, vol. 2 (1936), pp. 354–382.
5. F. Hausdorff, *Eine Ausdehnung des Parsevalschen Sätzes über Fourierreihen*, Math. Zeitschr., vol. 16 (1923), pp. 163–169.
6. F. Riesz, *Über die Randwerte einer analytischen Funktion*, Math. Zeitschr., vol. 18 (1923), pp. 87–95.

Peterhouse, Cambridge.

ACKNOWLEDGMENT

In bringing to a close the third volume of the Duke Mathematical Journal the editors wish to state that their task in launching the new periodical has been greatly lightened by the whole-hearted coöperation received from many sources: from Vice-President Flowers and other administrative officers of Duke University, who have provided the necessary funds in somewhat difficult times; from the mathematical public, which has aided by contributing articles, refereeing them and subscribing; and from the Waverly Press, which has cheerfully modified many of its practices to suit our needs. We are particularly indebted to Messrs. A. Cohen, J. D. Tamarkin and S. Lefschetz, who by arranging the transfer of papers from the established American journals helped us to publish our first number earlier than would otherwise have been possible; to Mr. R. E. Langer, who acted as editor for analysis during the academic year 1935–1936 when Mr. Widder was absent in Europe; to those mathematicians who, although not members of our board, have served as referees, namely,

J. W. Alexander,
H. Bateman,
P. Bernays,
S. Bochner,
D. G. Bourgin,
H. R. Brahana,
A. B. Brown,
H. B. Curry,
T. L. Downs,
W. B. Ford,
C. A. Garabedian,
V. G. Grove,
E. Hille,
E. Hopf,
V. A. Hoyle,
E. V. Huntington,
N. Jacobson,
F. E. Johnston,
S. Lefschetz,
N. Levinson,
D. C. Lewis, Jr.,
H. W. March,
M. Morse,
F. D. Murnaghan,
S. B. Myers,
W. C. Randels,
I. J. Schoenberg,
W. Seidel,
A. Sinkov,
N. E. Steenrod,
M. H. Stone,
G. Szegö,
J. D. Tamarkin,
H. S. Vandiver,
R. J. Walker,
M. J. Weiss,
H. Whitney,
O. Zariski,
L. Zippin,

and to Misses Hilda Howes and Alta Odoms, who have done the secretarial work and proof reading. To all of these we express our sincere thanks.

ournal
al has
many
Duke
times;
eferee-
erfully
debted
ng the
ublish
R. E.
6–1936
hough

g,

,

,

ctarial
s.